JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 114, D23301, doi:10.1029/2009JD012427, 2009

Validation of the Harvard Lyman-α in situ water vapor instrument: Implications for the mechanisms that control stratospheric water vapor

E. M. Weinstock,[1] J. B. Smith,[1] D. S. Sayres,[1] J. V. Pittman,[2] J. R. Spackman,[3,4] E. J. Hintsa,[4] T. F. Hanisco,[5] E. J. Moyer,[6] J. M. St. Clair,[7] M. R. Sargent,[1] and J. G. Anderson[1]

Received 5 May 2009; revised 17 August 2009; accepted 28 August 2009; published 2 December 2009.

[1] Building on previously published details of the laboratory calibrations of the Harvard Lyman-α photofragment fluorescence hygrometer (HWV) on the NASA ER-2 and WB-57 aircraft, we describe here the validation process for HWV, which includes laboratory calibrations and intercomparisons with other Harvard water vapor instruments at water vapor mixing ratios from 0 to 10 ppmv, followed by in-flight intercomparisons with the same Harvard hygrometers. The observed agreement exhibited in the laboratory and during intercomparisons helps corroborate the accuracy of HWV. In light of the validated accuracy of HWV, we present and evaluate a series of intercomparisons with satellite and balloon borne water vapor instruments made from the upper troposphere to the lower stratosphere in the tropics and midlatitudes. Whether on the NASA ER-2 or WB-57 aircraft, HWV has consistently measured about 1–1.5 ppmv higher than the balloon-borne NOAA/ESRL/GMD frost point hygrometer (CMDL), the NOAA Cryogenic Frost point Hygrometer (CFH), and the Microwave Limb Sounder (MLS) on the Aura satellite in regions of the atmosphere where water vapor is <10 ppmv. Comparisons in the tropics with the Halogen Occultation Experiment (HALOE) on the Upper Atmosphere Research Satellite show large variable differences near the tropopause that converge to ~10% above 460 K, with HWV higher. Results we show from the Aqua Validation and Intercomparison Experiment (AquaVIT) at the AIDA chamber in Karlsruhe do not reflect the observed in-flight differences. We illustrate that the interpretation of the results of comparisons between modeled and measured representations of the seasonal cycle of water entering the lower tropical stratosphere is dictated by which data set is used.

Citation: Weinstock, E. M., et al. (2009), Validation of the Harvard Lyman-α in situ water vapor instrument: Implications for the mechanisms that control stratospheric water vapor, *J. Geophys. Res.*, *114*, D23301, doi:10.1029/2009JD012427.

1. Introduction

[2] Water vapor is the dominant natural greenhouse gas. An accurate and continuous global water vapor measurement record in the upper troposphere and lower stratosphere (UT/LS) and the tropical tropopause layer (TTL) is pivotal for: (1) unraveling the relative importance of the dehydration mechanisms proposed to control the water vapor budget of the stratosphere; (2) determining the distribution of relative humidities within and in the vicinity of thin cirrus clouds; (3) quantifying the heterogeneous loss of stratospheric ozone in the Arctic and Antarctic both from water vapor's direct impact on the heterogeneous removal of ozone as well as its potential impact on vortex temperature; and (4) understanding the radiative properties of the TTL and stratosphere, especially at a time when the relationship between global climate change and surface and atmospheric temperatures must be clearly established.

[3] Since *Brewer* [1949] first postulated that the aridity of the stratosphere results from the dehydration of slowly rising air at the cold tropical tropopause, numerous attempts have been made to quantify the relationship between ice saturation at tropical tropopause temperatures and water vapor mixing ratios measured in the lower tropical stratosphere. Early attempts used an Eulerian approach, either

[1]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts, USA.
[2]Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts, USA.
[3]Chemical Sciences Division, Earth System Research Laboratory, NOAA, Boulder, Colorado, USA.
[4]Cooperative Institute for Research in Environmental Sciences, University of Colorado at Boulder, Boulder, Colorado, USA.
[5]Earth Sciences Division, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.
[6]Department of the Geophysical Sciences, University of Chicago, Chicago, Illinois, USA.
[7]Geology and Planetary Sciences Division, California Institute of Technology, Pasadena, California, USA.

0148-0227/09/2009JD012427$09.00

assuming preferential ascent in the western tropical Pacific [*Newell and Gould-Stewart*, 1981] or ascent in a full latitude circle with no longitudinal preference [*Mote et al.*, 1996; *Weinstock et al.*, 2001], or both possibilities [*Randel et al.*, 2004]. *Fueglistaler et al.* [2005] carried out Lagrangian back trajectory calculations in the tropics for the 1979–2001 period using reanalysis ERA-40 data from the European Center for Medium-range Weather Forecasting (ECMWF) to test the hypothesis that entry-level stratospheric water vapor is set by the saturation mixing ratio at the minimum temperature the parcel has experienced during the previous three months. The authors of these studies all made assumptions, both explicit and implicit. Two explicit assumptions were that convection contributed negligibly to the stratospheric water vapor budget and that air masses in the TTL are efficiently dehydrated to ice saturation corresponding to the lowest temperature experienced by the air parcel. The implicit assumption was that the accuracy of the selected water vapor data set is sufficient to constrain the mechanisms used to control humidity in the model being tested against the data. While reasoned arguments are often made to justify the explicit assumptions [e.g., *Fueglistaler et al.*, 2005], that is typically not the case for the implicit assumptions.

[4] There are, in fact, experimental data to question these explicit assumptions. Regarding convection, isotopic water vapor data reported by the ATMOS instrument [*Kuang et al.*, 2003] cannot be explained without convective injection into the TTL. *Dessler et al.* [2007] modified the *Fueglistaler et al.* [2005] trajectory model to include HDO as well as convective ice lofting in order explain stratospheric measurements of HDO and to examine the potential impact of convection on both H_2O and HDO. Their results show that convective ice lofting below the cold point can reproduce measured delta-D values by reducing HDO depletion with minimum impact on H_2O. However, the higher in the TTL that convection deposits ice and increases the water mixing ratio to local ice saturation, the less likely that the moistened air mass will be further dehydrated before entering the stratosphere. Regarding dehydration efficiency, there is a significant body of water vapor measurements in clear air and clouds in the TTL in which supersaturations of 30–60% predominate, with the higher supersaturations prevalent at temperatures below 190 K [e.g., *Jensen et al.*, 2005, 2007]. More recently, addressing these issues, *Read et al.* [2008] used the Microwave Limb Sounder (MLS) temperature data and a two-dimensional model that combined the Holton and Gettelman dehydration model [*Holton and Gettelman*, 2001] with a tropical convective model of *Folkins and Martin* [2005] to generate the seasonal cycle of water in the TTL to compare with MLS water vapor. Their dehydration model assumed condensation at 160% saturation over ice, and dehydration down to 100% saturation once a cloud is formed.

[5] Assessing the mechanisms that control stratospheric water vapor by comparing modeled and measured water vapor can be successful only if the stated uncertainties in water vapor data sets are in fact accurate. Unfortunately, ongoing significant systematic differences have been identified in upper tropospheric and lower stratospheric (UT/LS) water vapor measurements that extend well beyond stated uncertainties. Intercomparisons illustrating these differences have been compiled in chapter 2 of the 2000 SPARC Assessment of water vapor in the UT/LS [*Kley et al.*, 2000] and summarized in Figure 1 of that report. Kley et al.'s Figure 1 (bottom), which covers intercomparisons between 60 and 100 hPa, shows that the Harvard Lyman-α photofragment fluorescence hygrometer (HWV) and the JPL TDL hygrometer on the NASA ER-2 research aircraft measure about 10 to 50% higher than the instruments on other platforms. However, evaluation of the validation of the contributing instruments was beyond the scope of the report. Since publication of this report, systematic measurement differences continue to be observed between HWV on the NASA WB-57 research aircraft and the balloon-borne NOAA/ESRL/GMD frost point hygrometer (CMDL), CFH, the Halogen Occultation Experiment (HALOE), and MLS [*Read et al.*, 2007; *Vömel et al.*, 2007a]. Whatever approach is used to evaluate the importance of mechanisms hypothesized to control stratospheric humidity, without resolving water vapor measurement differences, the mechanistic mystery cannot be solved.

[6] Data taken during the Costa Rica Aura Validation Effort (CRAVE) in January and February 2006 provided a comprehensive set of water vapor intercomparison data in the lower stratosphere. These data consistently show 1.5–2 ppmv differences in the stratosphere between HWV and the Integrated Cavity Output Spectrometer (ICOS) [*Sayres et al.*, 2009], the Harvard in situ instruments on the WB-57, and both the NOAA Cryogenic Frost point Hygrometer (CFH) and MLS [e.g., *Read et al.*, 2007, Figure 24; *Vömel et al.*, 2007a]. These differences might be indicative of offsets, either positive in the Harvard instruments or negative in CFH and MLS. In response to these data, calibrations and intercomparisons were initiated in our laboratory to specifically explore instrument performance under low water conditions. Additionally, the summary report of a water vapor workshop organized as part of the CRAVE Science meeting directed instrument principal investigators to focus on laboratory calibrations and intercomparisons with water vapor mixing ratios of 0–5 ppmv.

[7] Although not explicitly stated, the overall thrust of the report is that the primary source of instrument validation must be laboratory calibrations traced to reliable standards. This is consistent with the philosophy that has constrained and guided the design and testing and of all the balloon and aircraft-borne instruments developed in our laboratory. We assert that a more stringent and necessary caveat to this is that the conditions under which laboratory calibrations are carried out must be representative of flight conditions. Otherwise, potential errors resulting from differences in laboratory and flight conditions must be addressed with the appropriate in-flight diagnostics. This validation process is not realistic for satellite borne instruments, which must rely on intercomparisons with other instruments over a wide range of water vapor mixing ratios, temperatures, and pressures for validation (e.g., for HALOE, *Harries et al.* [1996]; for MLS, *Read et al.* [2007]). The same validation process is followed by *Vömel et al.* [2007a], who quote 9% uncertainty for water vapor data in the tropopause region in their validation paper. Their quoted in-flight uncertainty is based solely on uncertainties in the measurement of the frost-point temperature and the ambient pressure measurement by the Vaisala RS80 radiosonde.

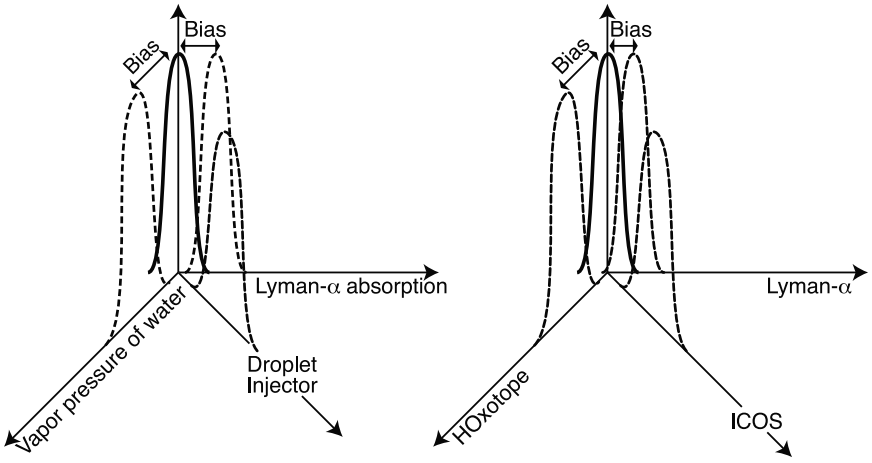

Figure 1. Representation of how (left) the use of independent calibration techniques and (right) the intercomparison in the laboratory and in flight of instruments with different detection techniques and sampling methods serve to minimize systematic errors in the flight measurement of water vapor.

[8] We describe in this manuscript progress in both laboratory calibrations and in-flight intercomparisons of in situ instruments on the WB-57 that constrain uncertainties in HWV or the Harvard Lyman-α photofragment fluorescence total water hygrometer (HTW) [*Weinstock et al.*, 2006a]. Based on that calibration accuracy and constraint, we assess the observed differences between water vapor measured by the balloon-borne CMDL and CFH frost point hygrometers, HALOE and MLS on the one hand and the aircraft-borne HWV and HTW instruments on the other. Special focus is placed on HWV, which has a long history of making stratospheric water vapor measurements on the NASA ER-2 and WB-57 research aircraft. We then assess the implications of water vapor measurement discrepancies on our ability to constrain the mechanisms that control stratospheric water vapor.

2. Lyman-α Water Vapor Instrument

2.1. Principle of Operation

[9] The principle of operation of HWV has been previously described in detail [*Weinstock et al.*, 1994]. Briefly, 121.6 nm (Lyman α) radiation from an RF discharge lamp photodissociates water vapor in air ram-fed into a 5.08 cm square duct at velocities of 40 to 100 m/s. A fraction of the resulting OH fragments are formed in their first excited electronic state ($A^2\Sigma^+$), and the resultant OH fluorescence at $\sim$315 nm is collected at right angles to the Lyman-α beam through a narrow-band filter and detected with a photomultiplier tube (PMT). In the upper troposphere and lower stratosphere, the observed detector signal is directly proportional to the water vapor volume mixing ratio. Lamp scatter near 315 nm, the source of the background counts detected by the PMT, is measured by using a quartz window that periodically blocks the Lyman-α beam. Changes in lamp intensity are monitored with a vacuum photodiode opposite the lamp and are used to normalize the fluorescence signal. Additionally a mirror is used to reflect a portion of the Lyman-α beam back across the tube to a second photodiode, providing the means to carry out radial absorption measurements in the laboratory and in flight when the water vapor concentration is sufficiently high, typically in the middle to upper troposphere.

[10] A detection module identical to the one described here is used in the HTW instrument. In contrast to the near ambient temperatures in the HWV detection axis, the air temperature in the HTW duct and detection axis, including its detection optics and components, are all near 290–295 K, because the HTW inlet is heated to evaporate ice particles in flight.

2.2. Instrument Accuracy

[11] We base the fundamental accuracy of HWV on laboratory calibrations that are referenced to fundamental physical constants of water and carried out under simulated flight conditions. We further establish that the calibration is maintained in flight, and perform diagnostics to identify and quantify in-flight sampling errors and/or instrument artifacts using in-flight diagnostics. Finally we make use of comparisons with other instruments, primarily other Harvard hygrometers. As we will illustrate shortly, a primary advantage of using other Harvard hygrometers in the validation process is that ongoing laboratory intercomparisons provide the opportunity to test for systematic errors under controlled sampling conditions. The results of these laboratory tests then provide a basis for evaluating comparable in-flight tests. We characterize potential errors as errors in calibration, sampling, or from instrument artifacts, and we establish instrument accuracy by determining or constraining these sources of error.

[12] The HWV instrument has had a quoted accuracy of 5% based on a continuous record of laboratory calibrations over the past 15 years, as well as in-flight diagnostics, in-flight validation using direct absorption, and intercomparison with the JPL tunable diode laser hygrometer [*Hintsa et al.*, 1999], as well as intercomparison with the HTW instrument [*Weinstock et al.*, 2006a, 2006b]. For the version of the Lyman-α instrument that flew on the ER-2, an instrument offset was constrained using laboratory measurements of the water vapor content of liquid nitrogen boil-off. The maximum offset was determined to be $\sim$<0.2 ppmv. This corresponds to a potential 7% offset during northern

Table 1. Calibration Techniques for Each Harvard Water Vapor Instrument

Instrument	Vapor Pressure of Liquid Water	Liquid Water Droplet Injector	Absorption at 121.6 nm
Lyman α[a]	Y	N	Y
ICOS[b]	Y	Y	N
HOxotope[b]	Y	Y	N

[a]These calibration methods are used for both HWV and HTW.
[b]Both ICOS (Integrated Cavity Absorption Spectrometer) and HOxotope (modified from the OH laser-induced fluorescence instrument) were developed at Harvard to measure water vapor isotopes. The droplet injector provides water droplets with a consistent and accurately measured diameter of about 70 microns at a controlled frequency [*St. Clair et al.*, 2008; *Sayres et al.*, 2009].

midlatitude winter in the TTL when water vapor mixing ratios are typically about 3 ppmv. In this section we focus on the results of recent laboratory calibrations, laboratory intercomparisons and in-flight intercomparisons to constrain uncertainties and instrument artifacts for the WB-57 HWV instrument at the low water vapor mixing ratios of the tropical upper troposphere and lower stratosphere.

2.3. Laboratory Calibration Methodology

[13] Each Harvard instrument conducts laboratory calibrations using pairs of measurable standards that are tied to a molecular or physical property of water and applicable to flight conditions. This pair-wise procedure, using the calibration techniques listed in Table 1, sets the fundamental accuracy of each instrument.

[14] Figure 1 (left) provides a schematic representation of the potential uncertainty associated with calibrating an instrument using two of the three independent reference standards, where the possible standards are identified by the axis labels, and each one is traceable to a molecular or physical property of water. The peak of the dashed Gaussian line shapes along each axis represents a potential bias associated with each calibration technique and its width represents the uncertainty associated with the precision of the technique. The solid Gaussian represents the minimization of measurement bias that is established by the required overlap of the instrument sensitivities derived from each of the independent calibration techniques. The better the overlap of the sensitivities derived using the individual calibration techniques and the better their respective precision, the smaller any potential systematic error in the final instrument calibration, and the narrower its uncertainty. Details of laboratory calibrations for the Harvard water instruments using pairs of the three calibration techniques have been thoroughly described in the literature [*Weinstock et al.*, 1994, 2006a; *St. Clair et al.*, 2008; *Sayres et al.*, 2009].

[15] Figure 1 (right) similarly illustrates how agreement in the laboratory (and likewise in flight) between the three Harvard water vapor instruments, each of which utilizes a different detection method and sampling technique, minimizes the potential bias in the measurements. Gaussians along each axis schematically represent the magnitude of potential biases and the associated precision of measurements made by each of the instruments. Measurement overlap for the three instruments constrains the magnitude of a bias, and instrument precision constrains the uncertainty in that bias. The validation of all three is established by the combined agreement during laboratory and in-flight comparisons. Agreement in the laboratory also confirms calibration consistency between the instruments over a wide range of water mixing ratios, including stratospheric values. Additionally, agreement in flight over a range of pressures and temperatures constrains the magnitude of potential artifacts not present in the laboratory.

2.4. In-Flight Measurement Validation Methodology

[16] For in situ water vapor instruments, potential systematic errors involve instrument and platform contamination, temperature or pressure dependence, water vapor mixing ratio dependence, and spurious instrument offsets (positive or negative). Contamination typically results from surfaces which exhale (or inhale) water for a period of time, resulting in hysteresis in the measurement. We list in Table 2 specific examples of how we use instrument B to validate the performance of instrument A as the indicated variable changes during the flight. While the ''comment'' column provides some information on the particulars of the intercomparison, we provide more details here for selected intercomparisons. The instruments compared in this study were HWV, HTW, HOxotope [*St. Clair et al.*, 2008], and ICOS [*Sayres et al.*, 2009]. The HTW instrument uses a detection module identical to that of HWV. However, as stated earlier, because it uses a heated inlet to evaporate ice particles in flight, the air temperature in the HTW duct as well as its detection optics and components are all near 290–295 K, as opposed to the near ambient temperatures in HWV. It samples air with an isokinetic inlet at flows of $\sim$6 m/s at the detection axis and typically requires a short drying out period during each flight [*Weinstock et al.*, 2006b]. HOxotope is a modified version of an OH flight instrument that used laser-induced fluorescence (LIF) to measure OH mixing ratios [*Wennberg et al.*, 1994]. The modification adds a module that uses an argon excimer lamp to produce ground electronic state OH from the UV photolysis of ambient water vapor, which is then measured by LIF. ICOS is a cavity absorption technique that uses a continuous wave (cw) infrared tunable diode laser to obtain near-simultaneous measurements of H_2O, HDO, and $H_2^{18}O$. Light from a high-power continuous wave laser is injected into a high-finesse optical cavity consisting of a pair of highly reflective mirrors (R $\approx$ .09998) and containing the atmospheric gas to be measured. Because light is trapped within the optical cell for tens of microseconds, however, the effective optical path length is several kilometers rather than the tens of meters of a typical Herriott cell, and the instrument sensitivity is correspondingly greater. Both HOxotope and ICOS have flow rates of 5–10 m/s, and require some drying out period at the beginning of each flight. The instruments fly mounted in WB-57 pallets, and their inlets extend well beyond the aircraft boundary layer. Nevertheless, during their first mission, the Aura Validation Experiment–Water Isotope Intercomparison Flights (AVE-WIIF), ICOS and HOxotope exhibited comparable levels of contamination during the first hour of each flight, as evidenced by intercomparison with HWV. This was diagnosed as contamination from the aircraft fuselage as opposed to instrument contamination because of the markedly different flush times for these two instruments

Table 2. Illustrated Sensitivity of Instrument A to a Specific Variable by Comparison With Instrument B Over the Range of That Variable in the TTL

Variable	Instrument A	Instrument B	Comment
Flow velocity and/or rapid or large changes in H_2O mixing ratio	HOxotope, ICOS, HTW	HWV	Test for sampling problems manifested as instrument hysteresis during flight segments where water vapor mixing ratio undergoes rapid order of magnitude changes.
Temperature	HWV	HTW	The heated inlet of the total water instrument maintains the temperature at the detection axis close to 293 K.
Pressure, temperature	HWV	HOxotope, ICOS	HOxotope and ICOS maintain their cell pressures at 30 hPa and ~ 295 K.
H_2O mixing ratio	HWV	ICOS	Comparison of a fluorescence technique with an absorption technique checks linearity of the two instruments from a few to hundreds of ppmv.

[*St. Clair et al.*, 2008]. Intercomparison with HWV during these flights has been used as an indicator of how long after aircraft takeoff hysteresis becomes negligible for HTW, ICOS and HOxotope.

[17] The first row in Table 2 lists the use of HWV to test for instrument hysteresis in the other instruments. HWV, because of its forward facing, ram-fed inlet providing flows of ~70 m/s at the detection axis, has been shown to sample air uncontaminated by either the aircraft fuselage or the instrument walls [*Hintsa et al.*, 1999; *Weinstock et al.*, 2006b]. Continuing with the next intercomparison, we illustrate the insensitivity of HWV's calibration to temperature by comparing HWV with HTW. The agreement in clear air of HWV with HTW, over a range of temperatures, (e.g., see Figures 4 and 5 below) validates that Lyman-α detection axis sensitivities derived from laboratory calibrations at room temperature and atmospheric pressures are valid under flight sampling conditions.

[18] The next row shows that in-flight intercomparisons between HWV and ICOS or HOxotope can be used to validate the laboratory determination of the pressure dependence of the HWV calibration. The detection axes of both the ICOS and HOxotope instruments operate at the same constant pressure in flight and in the laboratory and thus provide a stable reference with respect to pressure changes in flight. Furthermore, their detection regions also operate under constant temperature conditions, so comparisons with HOxotope and ICOS also serve to validate the insensitivity of the HWV detection axis to temperature.

[19] The final intercomparison between HWV and ICOS over the wide range of ambient water vapor mixing ratios confirms the linearity of the sensitivity of both instruments to the water vapor mixing ratio.

2.5. Low Water Calibrations and Laboratory Intercomparisons

[20] A detailed description of the HWV calibration procedure has been published [e.g., *Weinstock et al.*, 2006a], thus we provide only a brief summary here. The laboratory calibrations are carried out by measuring the OH fluorescence signal resulting from Lyman-α excitation of humidified air over a range of air densities and water vapor mixing ratios. The specific humidity of the air is controlled by the stepwise addition of a slow flow of saturated air to a primary flow of dry air. The secondary flow is brought to saturation by bubbling it through water, using a two-stage apparatus. The slope of a plot of OH fluorescence counts/s versus the water vapor mixing ratio of the airflow determines the sensitivity of the detection axis to ppmv of water vapor at a given air density. The intercept is a measure of residual water measured by the instrument, resulting from water vapor in the dry air, from instrument surfaces, or from any instrument artifact that is a source of fluorescence signal proportional to Lyman-α flux. The calibration accuracy for HWV using this method is ±5%. Here we focus on laboratory calibrations and tests run with water vapor mixing ratios between 2 and 10 ppmv, comparable to the values encountered in the TTL. Because typical laboratory calibrations are carried out at room temperature, their insensitivity to temperature must also be established. Additionally, when carrying out titrations at these humidities, the difference between measured water vapor and that added using the "bubbler" can vary depending on the choice of carrier gas (dry nitrogen or air) and the degree to which the calibration system has been purged with dry nitrogen and the actual humidity of the main carrier gas used in the titration.

[21] In each plot of Figure 2 we show the stepwise addition of small increments of water vapor as determined by the vapor pressure of water over a bubbler. The water added through the bubbler + 1.5 ppmv is plotted in red. The addition of 1.5 ppmv accounts for the sum of water vapor in the carrier gas and residual water vapor in the instrument. The measured water mixing ratio, determined using the sensitivity of the axis from previous calibration runs, is plotted in blue. Throughout the entire range of temperatures covered in this run, measured water vapor agrees with the sum of bubbler water vapor +1.5 ppmv to within ±0.1 ppmv. Together these plots not only illustrate that the detection axis sensitivity is independent of temperature, but also that any potential offset is insensitive to temperature because the residual water vapor in the dry air is independent of temperature, at least down to 250 K. Though this is somewhat higher than ambient temperatures seen in the TTL, because of ram heating in the duct, the actual temperatures at the detection axis are ~15 degrees above ambient, and thus closer to the minimum temperature reached in these laboratory tests. For example, air temperatures seen during the 2007 Tropical Composition, Cloud and Climate Coupling (TC4) mission over Costa Rica ranged from 198 to 205 K for the lowest water vapor mixing ratios, while the corresponding air temperatures in

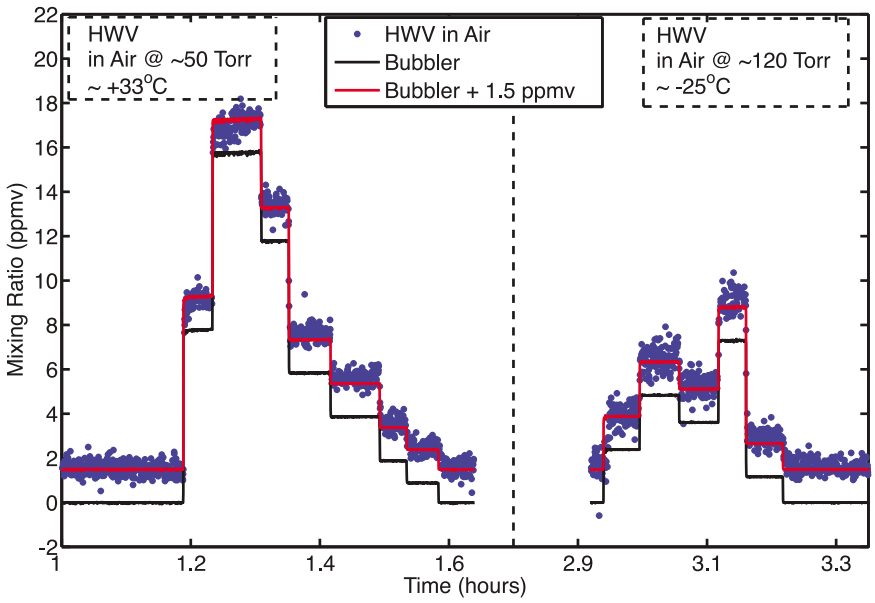

Figure 2. Repeat laboratory calibrations at temperatures ~60°C apart illustrate temperature independence. Data in both plots are analyzed using the same calibration constants.

the water vapor duct were between 213 and 220 K. This is important because it provides independent evidence that room temperature offset constraints determined from laboratory measurements are valid under flight measurement conditions.

[22] Figure 3 shows representative low water vapor calibrations of both the ER-2 (Figure 3, left) and WB-57 (Figure 3, right) detection axes, respectively. In order to focus on low water calibrations that require the driest possible carrier gas to constrain the magnitude of a possible artifact, these calibrations were carried out using nitrogen boiloff from a liquid nitrogen dewar. We include laboratory data from the ER-2 axis because some of the flight intercomparison data we show later in the manuscript include data from that instrument. The ICOS instrument was set up in parallel to the WB-57 axis and its measurement of water vapor is included. Its sensitivity was determined from previous calibrations. It was configured for use with a laser line coinciding with a very strong water vapor absorption, thus yielding a statistical uncertainty in water vapor measured by the ICOS instrument of about 0.01 ppmv. Configured in this way, ICOS is the ideal laboratory instrument for accurately and precisely measuring fractional ppmv water vapor mixing ratios. Uncertainties in the sensitivity from calibrations or absorption line strength parameters produce negligibly small absolute uncertainties in the measured mixing ratio. In contrast to HWV, the only instrument offset to which ICOS is susceptible is water outgassing from instrument walls. The outstanding signal-to-noise of ICOS provides the precision to observe the rapid time response of the ICOS instrument to sub-ppmv changes in added water vapor, and the data show no evidence of a measurable wall effect. We conclude, therefore, that the 0.2 ppmv offset observed by both ICOS and the Lyman-α instrument when the bubbler flow is equal to zero is most likely due to residual water vapor in the primary carrier flow. These low water vapor runs in the laboratory effectively constrain the magnitude of measurement artifacts

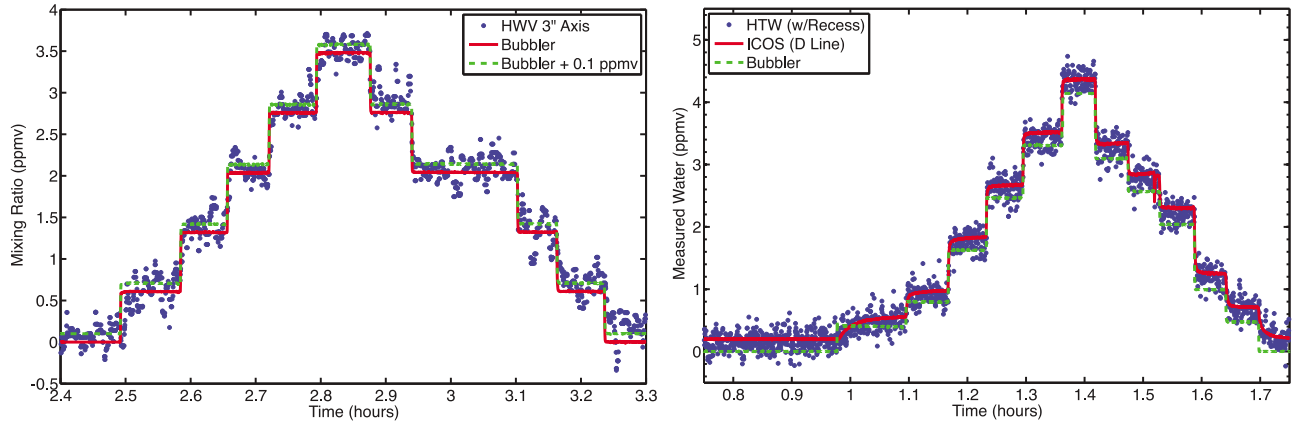

Figure 3. Low water vapor calibration runs for the (left) ER-2 and (right) WB57 total water detection axes.

Table 3. NASA Aircraft Missions That Provide Intercomparison Data for This Study With the Instruments Included in the Given Intercomparisons[a]

Mission	Aircraft	Location	Date	In Situ Instruments	Instruments Intercompared With
CWVCS	NASA WB-57	San Jose, Costa Rica	August 2001	HWV HTW	
AVE-WIIF	NASA WB-57	Houston, TX	July 2006	HWV ALIAS[b] ICOS HOxotope HTW	
CRAVE	NASA WB-57	San Jose, Costa Rica	February 2006	HWV ICOS	CFH MLS
TC4	NASA WB-57	San Jose, Costa Rica	August 2007	HWV HOxotope ICOS	CFH MLS
AVE test flights	NASA WB-57	Houston, TX	November 2004	HWV	CFH
Pre-AVE	NASA WB-57	San Jose, Costa Rica	January 2004	HWV	CMDL HALOE
CEPEX	NASA ER-2	Fiji	March 1993 to April 1993	HWV	CMDL, HALOE
SPADE	NASA ER-2	Moffett Field, CA	May 1993	HWV	CMDL
STRAT	NASA ER-2	Honolulu, HI	November 1995 to December 1996	HWV	HALOE
POLARIS	NASA ER-2	Honolulu, HI	September 1997	HWV	HALOE

[a]Missions: CWVCS, Clouds and Water vapor in the Climate System; AVE-WIIF, Aura Validation Experiment−Water Isotope Intercomparison Flights; CRAVE, Costa Rica Aura Validation Experiment; TC4, Tropical Composition, Cloud and Climate Coupling Experiment; AVE (test flights), Aura Validation Experiment; Pre-AVE, Pre-Aura Validation Experiment; CEPEX, Central Equatorial Pacific Experiment; SPADE, Stratospheric Photochemistry, Aerosols and Dynamics Expedition; STRAT, Stratospheric Tracers of Atmospheric Transport; Polaris, Photochemistry of Ozone Loss in the Arctic Region in Summer.
[b]ALIAS is Atmospheric Laser Infrared Absorption Spectrometer (NASA JPL).

affecting the Lyman-α instruments to $\sim$<0.1 ppmv in nitrogen and <0.2 ppmv in air.

3. Instrument In-Flight Intercomparison

[23] We now turn to in-flight intercomparisons that help validate instrument performance in flight. We list in Table 3 those missions that provide intercomparison data for this study for HWV since being developed for use on the NASA ER-2 research aircraft. Listed as well are the other in situ water vapor measurements used for comparison, whether made by instruments on the aircraft, or on a balloon or satellite acquiring data near the aircraft flight track. We first analyze data from those missions that provide the opportunity to intercompare Harvard instruments and to test instrument sensitivities to the variables in Table 2. Then we proceed to summarize the results of numerous intercomparisons between HWV and various satellite and balloon instruments over the past decade and a half.

[24] The AVE-WIIF mission provided the opportunity for an intercomparison of not only four Harvard instruments, but MLS as well. We therefore follow with a summary of intercomparisons between HWV and MLS. Because CFH flew during the MLS validation missions, we transition next to intercomparisons with CFH. We follow with intercomparisons with CMDL. CFH is an improved version of the CMDL instrument and intercomparison results [*Vömel et al.*, 2007a] show that the two instruments agree within their stated uncertainties. Accordingly, intercomparisons between HWV and both CFH and CMDL should be self-consistent. Finally, with the Pre-AVE campaign, we introduce intercomparisons with HALOE.

3.1. CWVCS Intercomparisons

[25] Focusing on intercomparisons in the TTL, we first show an intercomparison between HWV and HTW from the CWVCS mission. We use the results of this intercomparison to confirm the laboratory data plotted in Figure 2 showing that the sensitivity of HWV is independent of temperature. Figure 4 (top) plots potential temperature versus time for the flight on 20010809, illustrating multiple vertical transects through the tropopause region in cloud-free air with ambient temperatures ranging from 195 to 208 K. We plot in Figure 4 (middle) water vapor and total water versus time, with the difference between the two in Figure 4 (bottom). After the initial drying out of the total water inlet and duct, the average difference between HTW and HWV is 0.03 ppmv with a standard deviation of ±0.26 ppmv for the entire portion of the flight above 360 K, thus providing clear and conclusive evidence of the insensitivity of the HWV axis to temperature. Moreover, this intercomparison illustrates how HWV can pinpoint when hysteresis in the total water instrument becomes negligible. Aircraft takeoff was at 46630 s UT, and agreement between the two instruments was better than 5% in the stratosphere at 47280 s, 38 min after takeoff and 27 min after normal flow through HTW started. This level of agreement in clear air serves to validate the in-flight performance of each instrument and places an upper bound to the potential systematic errors to which each may be subject.

3.2. AVE-WIIF Intercomparisons

[26] The AVE-WIIF mission provided a unique opportunity for a multi-instrument intercomparison on the same platform, with six independent measurements of water vapor available, including four from our laboratory. Table 4 lists the six instruments, their detection techniques and sampling methods. While the focus here is on intercomparisons among the Harvard instruments, ALIAS data are also included. The NASA Jet Propulsion laboratory Laser Hygrometer (JLH) are not (R. Herman, personal communication, 2009). Agreement from the upper troposphere through the lower stratosphere covered almost two orders of magnitude in water vapor and is especially significant because the data were taken in a blind intercomparison, with no principal investigator (PI) seeing data from another instrument prior to submission, with the exception of HWV and HTW whose data are reported by a single PI.

[27] Agreement between the Harvard instruments was in general very good throughout the mission except during the ascent portions of the flight and during regions of poor laser performance or alignment [*St. Clair et al.*, 2008]. We show

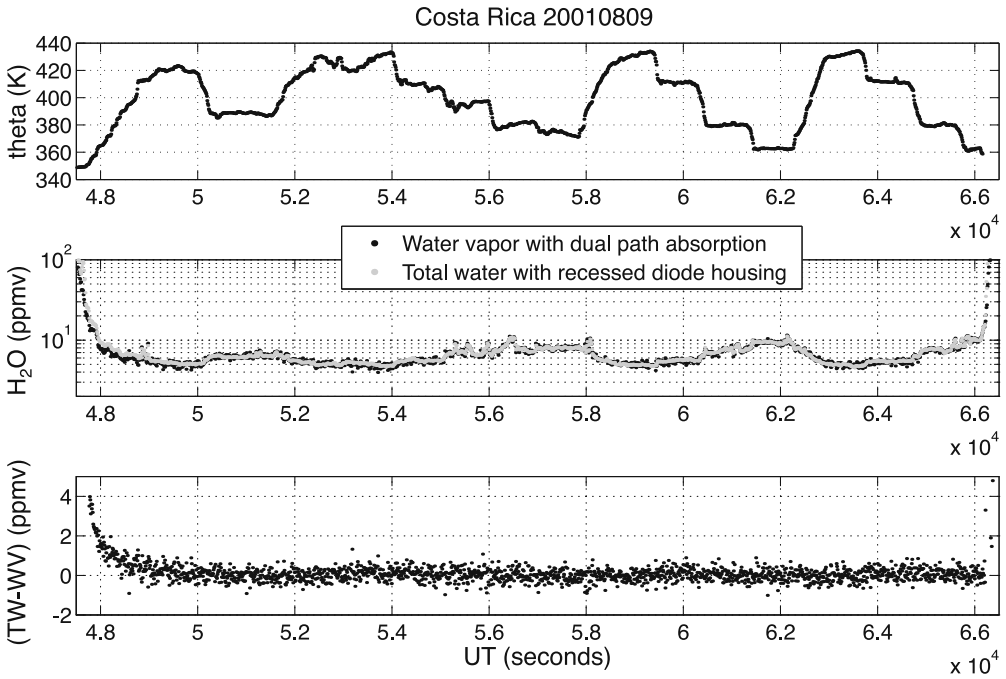

Figure 4. Transects of the TTL during flights from Costa Rica on August 9, 2001, showing (top) potential temperature, (middle) water vapor, and (bottom) difference between the total water and water vapor measurements versus time.

in Figure 5 (top) a segment of the 20050707 flight that transitions from the troposphere to the stratosphere, with differences of each instrument from the average mixing ratio plotted in Figure 5 (bottom). The full segment shows excellent agreement during the transition from 20 ppmv down to 5 ppmv while the ambient pressure changes from 190 to about 80 hPa and the temperature decreases from 218 to 195 K. The large stars, arbitrarily placed along the abscissa of Figure 5 (bottom), represent the average water vapor mixing ratio difference for each instrument from the mean along this flight segment, and are all less than ±0.05 ppmv.

[28] The agreement between the five in situ instruments in the stratosphere, exhibited during the 7 July AVE-WIIF flight and illustrated in Figure 5, is generally better than 5% or about 0.25 ppmv, as can be seen from the average differences represented by the color-coded stars. We believe that the gradual increase in the value of delta HWV/average through the plotted segment results from slight hysteresis in the slower flow hygrometers. The agreement exhibited at these low mixing ratios provides evidence that once instrument (internal) and aircraft (external) contamination has dissipated, the instruments agree over a wide range of temperatures and pressures. We reference Table 2 as a guide for how these data further validate the in-flight accuracy of the Harvard hygrometers. During this mission, four different Harvard instruments use three completely independent detection methods and four different sampling strategies and agree over a range of atmospheric pressures, temperatures, water vapor concentrations, and flow velocities. Combined with extensive laboratory calibration data, and laboratory intercomparisons as illustrated in Figure 3, this serves as an in-flight validation of the Harvard instruments.

[29] Least squares fits to the HOxotope, ICOS, and HWV data are shown in Figure 6. These plots exclude ICOS and HOxotope ascent data, where instrument hysteresis was a recognized problem, apparently caused mostly by contamination from the aircraft fuselage [*St. Clair et al.*, 2008]. Figure 5 illustrates, per row one of Table 2, that for the selected flight segment, HWV validates that the slower flow Harvard hygrometers are free from hysteresis. The data plotted in Figure 6 that show agreement among the three independently calibrated water vapor instruments confirm, per row three of Table 2, that the calibration of HWV is valid over the full range of atmospheric temperatures and pressures. The agreement of HWV and ICOS, per row four of the table, confirms the linearity of the calibration of both an absorption and a fluorescence technique from about 5 ppmv to 100 ppmv. Figures 5 and 6 also illustrate that any bias greater than about 0.25 ppmv in HWV and HTW

Table 4. Water Instruments on the WB-57 Aircraft During AVE-WIIF

Instrument	Detection Technique	Sampling Method
HWV	Photofragment fluorescence	Fast flow (20–100 m/sec) double-ducted forward facing inlet
HTW	Photofragment fluorescence	Heated forward-facing isokinetic inlet
ICOS	Infrared cavity output spectrometer	Pressure- and temperature-controlled rear-facing inlet
HOxotope	Photofragment laser-induced fluorescence	Pressure- and temperature-controlled, 8.3 l/s flow rate, rear-facing inlet
ALIAS	Infrared multipass tunable diode laser absorption	Heated forward-facing isokinetic inlet

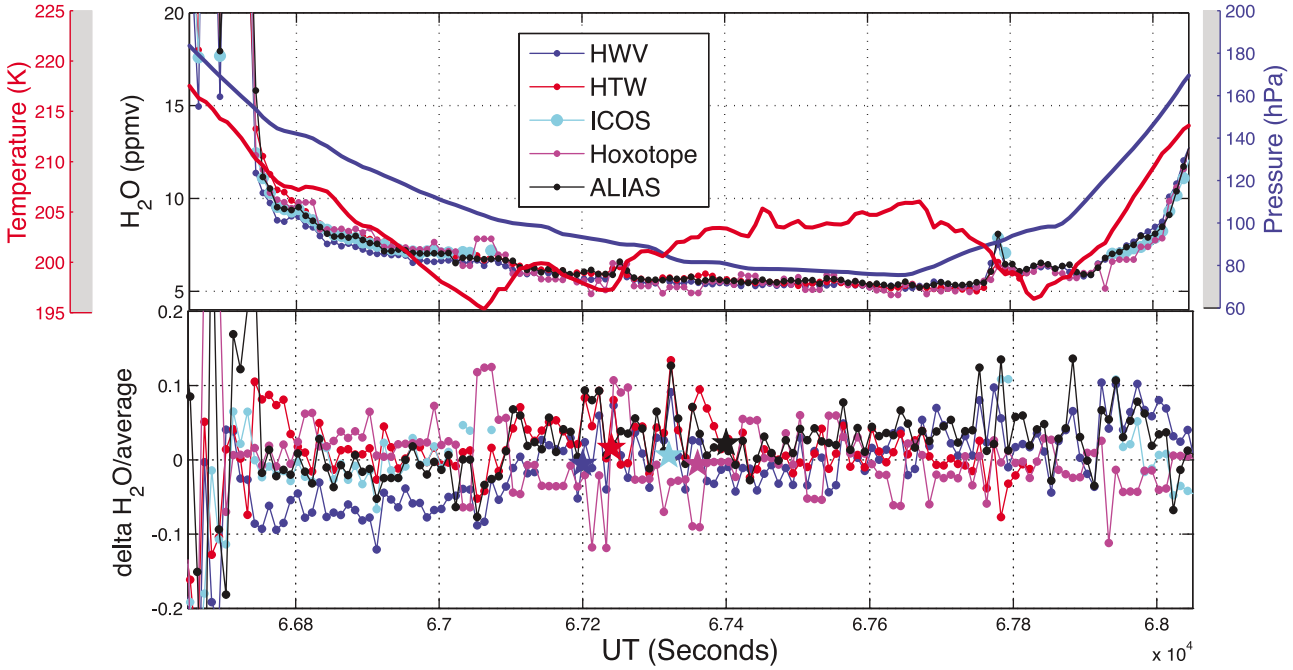

Figure 5. (top) Ambient pressure (dashed blue line) and temperature (dashed red line) are plotted along with water vapor data taken during the last stratospheric segment of the 20050707 flight. (bottom) Fractional differences from the average measured water vapor for each measurement for the same segment. The average water value is determined without ICOS because ICOS does not report data for a significant portion of the plotted segment. Also plotted as stars are the fractional differences averaged over this segment of the flight color-coded by instrument limited to data with average water values < 10 ppmv.

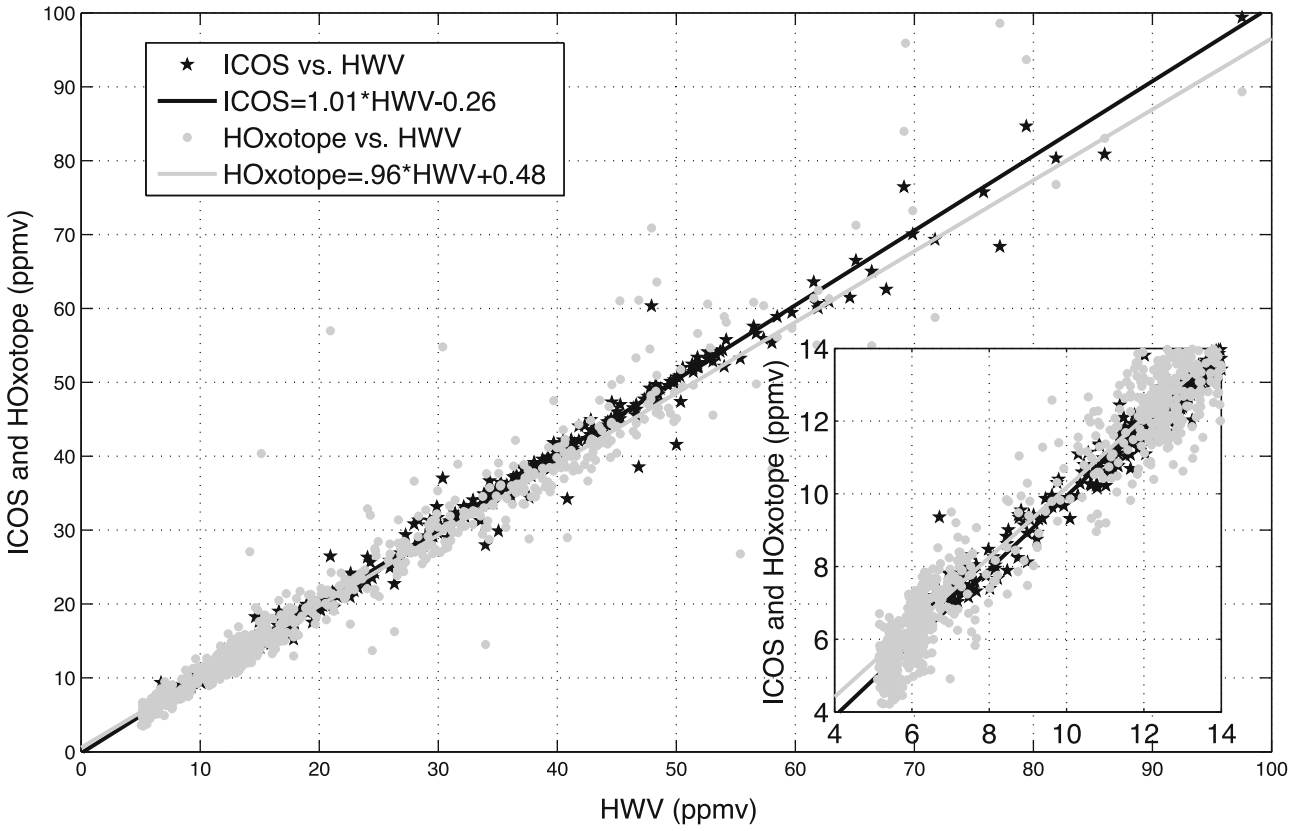

Figure 6. Least squares fits to the raw data for the last AVE-WIIF flight where HOxotope and ICOS are each plotted against HWV.

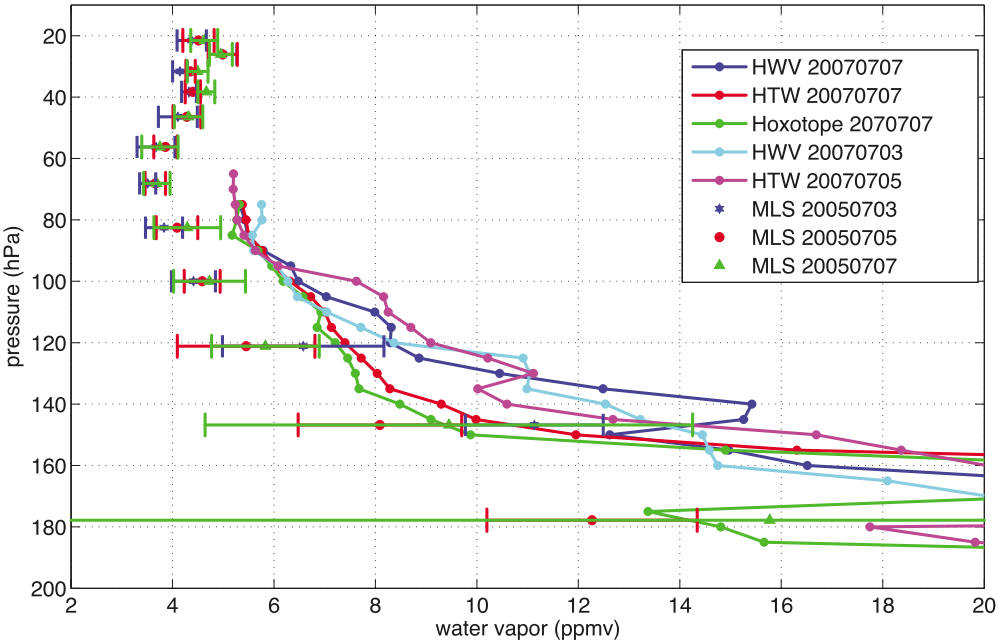

Figure 7. Intercomparison between water vapor measured on the WB57 during the three AVE-WIIF intercomparison flights and MLS water vapor profiles taken on the same days as the flights and with location constraints described in the text. The MLS vertical resolution in this region is 3.5 km [*Read et al.*, 2007].

data would also have to exist in both ICOS and HOxotope data.

[30] It is at low mixing ratios where such a bias is a concern because it is precisely at these low mixing ratios that the largest percentage disagreement exists with the frost point hygrometers and MLS. The first example of this disagreement is shown in Figure 7. MLS version 2.2 data taken on the three AVE-WIIF science flight days are plotted along with the in situ profiles as described in Figure 7's legend and caption. Version 2.2 MLS provide data at discreet pressure levels through the troposphere and stratosphere. The MLS data are presented as averages and standard deviations of all the profiles for the given date bounded by 20° and 35°N in latitude and 80° and 100°W in longitude. We purposely relax overlap requirements with the WB-57 flight path because the in situ profiles suggest little variability for stratospheric water vapor during this period of time. The standard deviations in the MLS data, color-coded by date, support this decision. Note, however, that atmospheric variability becomes significant in the upper troposphere. The spread between the average profiles of HOxotope and HWV in the upper troposphere on the flight of 7 July is a result of differences in the region sampled. The HWV profile represents an average of data taken over the entire flight, whereas the HOxotope profile represents an average over the last upper tropospheric and stratospheric segment.

[31] Based on the highest pressure level reached by the WB-57 on the flight of 5 July, we can quantify measurement differences between HWV and MLS at the 68, 82.5, and 100 hPa MLS pressure levels in the stratosphere. We also include intercomparison results from the 120 hPa MLS level in the upper troposphere. The average differences at the four pressure levels starting at 68 hPa are 1.6, 1.5, 1.8, and 2.6 ppmv. As expected, atmospheric variability appears to be increasing with increasing pressure, and on 5 July the in situ data at 100 hPa suggest the in situ sampling of a different air mass than the MLS instrument. At 120 hPa atmospheric variability is clearly increasing as evidenced by the spread in the in situ data, and the larger average difference of 2.6 ppmv is likely indicative of that variability. Accordingly, we omit the data at 120 hPa and pressure levels in the troposphere used in our estimate of the average difference between HWV and MLS, which is 1.6 ± 0.1 ppmv. The accuracy of MLS water vapor data in the UT/LS is based on agreement with CFH. It is quoted as 2.3 ± 11.8% at 68 hPa and below, and 6.4 ± 22% at 100 hPa [*Vömel et al.*, 2007a]. Accordingly, MLS acts as a surrogate for CFH in Figure 7.

3.3. CRAVE and TC4 Intercomparisons

[32] We turn next to the CRAVE and TC4 missions, which provide by far the majority of intercomparison opportunities between HWV, CFH, and MLS. A primary objective for both the CRAVE and TC4 missions was to validate instruments on the Aura satellite. For water vapor, this included CFH launches from Costa Rica during the CRAVE mission and from both Costa Rica and the Galapagos Islands during TC4. The timing of the Aura satellite overpass near Costa Rica and weather restrictions on the WB-57 launch window during TC4 prevented coordination of the WB-57 flight path with Aura overpasses. We therefore include MLS profiles between 12°N and 2°S latitude, and 75°W and 90°W longitude, the area that approximately overlaps the flight paths of the WB-57 during the mission, and further constrain the MLS profiles to be on the same day as the flights. For the CRAVE mission, we include MLS profiles between 10°N and 0°S latitude, and 75°W and 88°W longitude from 20060201 to

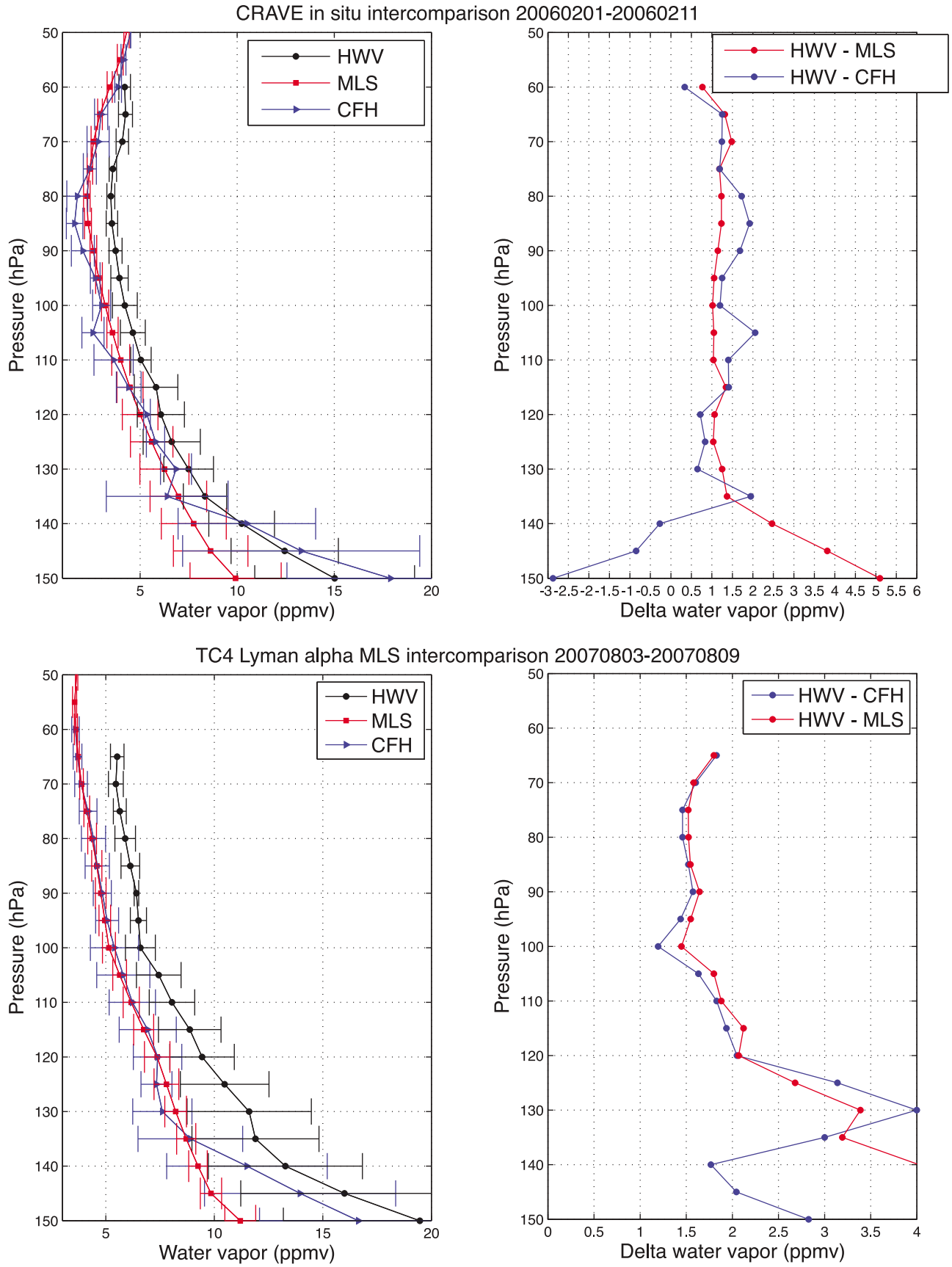

Figure 8. (top) Results of intercomparing HWV with MLS and CFH during the in situ portion of the CRAVE campaign, where the data are all binned and average at 5 hPa pressure intervals. (bottom) Similar data for the TC4 mission.

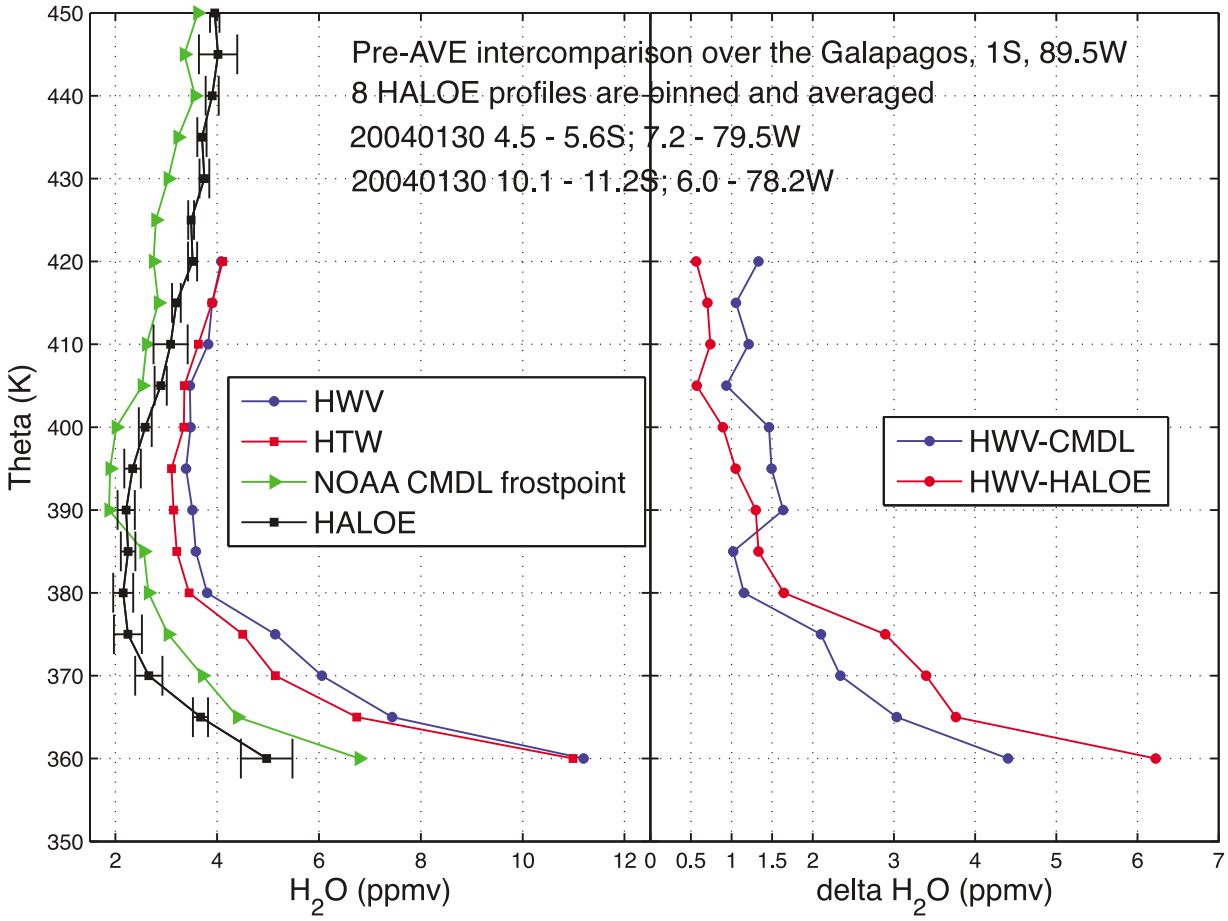

Figure 9. (left) Intercomparison of Harvard water vapor instruments on the NASA WB-57 research aircraft with HALOE and CMDL. CMDL data were taken on ascent. Data from HTW are included because all the measurements are in cloud-free air. HALOE data are corrected according to E. Remsberg (personal communication, 2005) and consistent with retrieval sensitivity studies described in section 2.3.1 of the SPARC 2000 report [*Kley et al.*, 2000], with increases ranging from a maximum of 1.2 ppmv near the tropopause to 1.05 ppmv at pressures below 56 hPa. (right) Differences between HWV and CMDL and HALOE.

20060211. For the CFH profiles, we only use sonde data taken on WB-57 flight days, and choose the sonde location, launched from Costa Rica or the Galapagos, depending on the WB-57 flight path through the stratosphere. While both ascent and descent data are archived for each sonde, we use the higher resolution ascent data where available (H. Vömel, personal communication, 2008). In Figure 8 we show intercomparisons between HWV, CFH and MLS for CRAVE and TC4. The plotted data represent the average profile over the entire mission. For the in situ instruments, the data are binned and averaged at 5 hPa intervals, and the horizontal error bars are the 1-sigma standard deviations based solely on the statistical variation about the mean. For MLS, the profiles on each date are interpolated and then binned and averaged at 5 hPa intervals. These intercomparisons illustrate that the bias of about 1.5 ppmv between HWV and both CFH and MLS is consistent for these two missions, and also consistent with the AVE-WIIF results. While there is no a priori reason to plot absolute as opposed to fractional differences between these instruments, there is less consistency in the observed fractional differences, especially when comparing CFH and HWV for the two missions.

3.4. AVE Test Flight Intercomparison

[33] On 18 November 2004, as part of the AVE test flight series, an intercomparison was arranged between CFH launched from Midland, TX, and HWV on the WB-57 aircraft. The aircraft performed a spiral profile around the balloon both while it was ascending and descending. The data in this intercomparison, which also include data from an aircraft-borne frost point hygrometer, are shown in Figure 13 of *Vömel et al.* [2007b]. Briefly, there is a systematic difference between HWV and CFH of about 3.5 ppmv in the lower stratosphere, which is about 2 ppmv greater than typically observed. One could speculate whether this difference is caused by an undiagnosed artifact of one of the instruments or atmospheric variability. Vömel et al. note that the data from an aircraft frost point hygrometer agree better with the balloon frost point hygrometer than with our instrument. From this, Vömel et al. suggest that the differences may be related to instrument technique or to the implementation of that technique, rather than differences

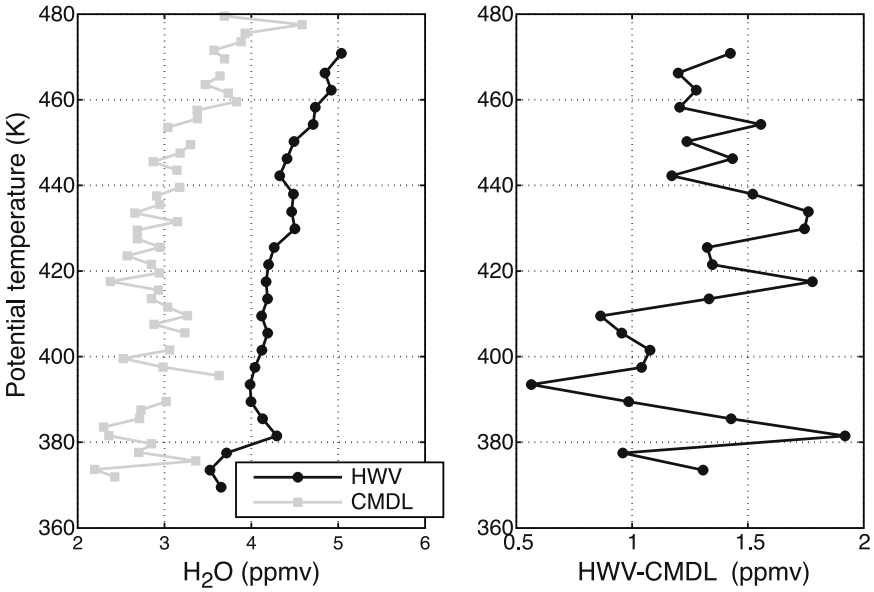

Figure 10. Intercomparison between HWV on the ER-2 and CMDL during CEPEX. Harvard points are binned and averaged at 2 K intervals for data taken during aircraft dives at 2°S latitude on 18, 21, 24, 29, and 31 March and 4 April 2003. Similarly binned and averaged CMDL data were taken on 12, 13, 14, 15, and 17 March, from launches off the U.S. *Vickers* research vessel traveling eastward at 2°S latitude, and from Christmas Island (2°N, 157°W) on 22 and 24 March.

due to platform. A more discerning look at *Vömel et al.*'s [2007b] Figure 13 reveals that the two aircraft water vapor profiles are similar in shape and differ from that of the balloon frost point profile. From this perspective, the observed differences could be exacerbated by air mass differences, which decrease at the highest aircraft altitudes. In fact, AVE WB-57 flights on the 7, 9, and 12 November all show extreme variability in ozone in the lower stratosphere. Agreement with HTW during the flights following the 18 November flight suggests no unusual behavior in water vapor. In any event, while the differences observed in this intercomparison are significantly different from the typical difference between the two instruments of about 1.5 ppmv, there is no way to draw any conclusions regarding instrument accuracy or performance from a single anomalous intercomparison.

3.5. Pre-AVE Intercomparisons

[34] We now turn to earlier missions to evaluate the consistency of the differences measured between HWV and CMDL. In preparation for validation of instruments on the Aura satellite, intercomparison experiments were carried out with the HALOE water vapor instrument, CMDL, and several in situ water vapor instruments on the NASA WB-57 research aircraft. In Figure 9 we illustrate the results of such an intercomparison that took place over Costa Rica during the Pre-AVE mission in January 2004. For the intercomparison, CMDL was launched from the Galapagos Islands (1°S, 89.5°W) to provide a water vapor profile coinciding with an overpass of the Upper Atmosphere Research Satellite (UARS) that carried the HALOE instrument. The WB-57 performed a spiral ascent around the balloon as it ascended through the TTL. The HALOE water vapor data shown here and throughout the paper are version 19 data with the pressure dependent correction ranging from a maximum of 1.2 ppmv near the tropopause to 1.05 ppmv at pressures below 56 hPa (E. Remsberg, personal communication, 2005; section 2.3.1 of the SPARC 2000 report [*Kley et al.*, 2000]). These data have an altitude resolution of approximately 2 km.

[35] The differences exhibited between the Harvard instruments and CMDL are consistent with the previously described intercomparisons. The disagreement in the stratosphere between HWV and HALOE gradually decreases from about 1.5 ppmv at 390 K to 0.5 ppmv at 420 K. These limited results above the tropopause region suggest that if a 0.5 ppmv difference between HALOE and HWV is representative of the stratosphere then these instruments are in agreement in the stratosphere within their uncertainties. Considering the ~20% high bias of HWV compared to HALOE shown in Figure 1 of the SPARC report, this result is encouraging, and more will be said about this in section 3.8 when intercomparisons using STRAT and POLARIS tropical data are discussed.

3.6. CEPEX Intercomparisons

[36] We next focus on intercomparisons with HWV on the NASA ER-2 [*Hintsa et al.*, 1999] and CMDL. In March 1993, both HWV and CMDL participated in the Central Equatorial Pacific Experiment (CEPEX). The ER-2 flight paths typically included stratospheric segments with profiles at 2°S, while CMDL data were taken on ascent on balloons launched from the *Vickers* research vessel, sailing eastward on the 2°S parallel of latitude from 165°E to 157°W longitude, and from Christmas Island (2°N, 157°E).

[37] Figure 10 compares binned and averaged CMDL data taken between 12 and 24 March, with binned and averaged HWV data taken from the NASA ER-2 during dives at 2°S on 19930318, 19930321, 19930324, 19930329,

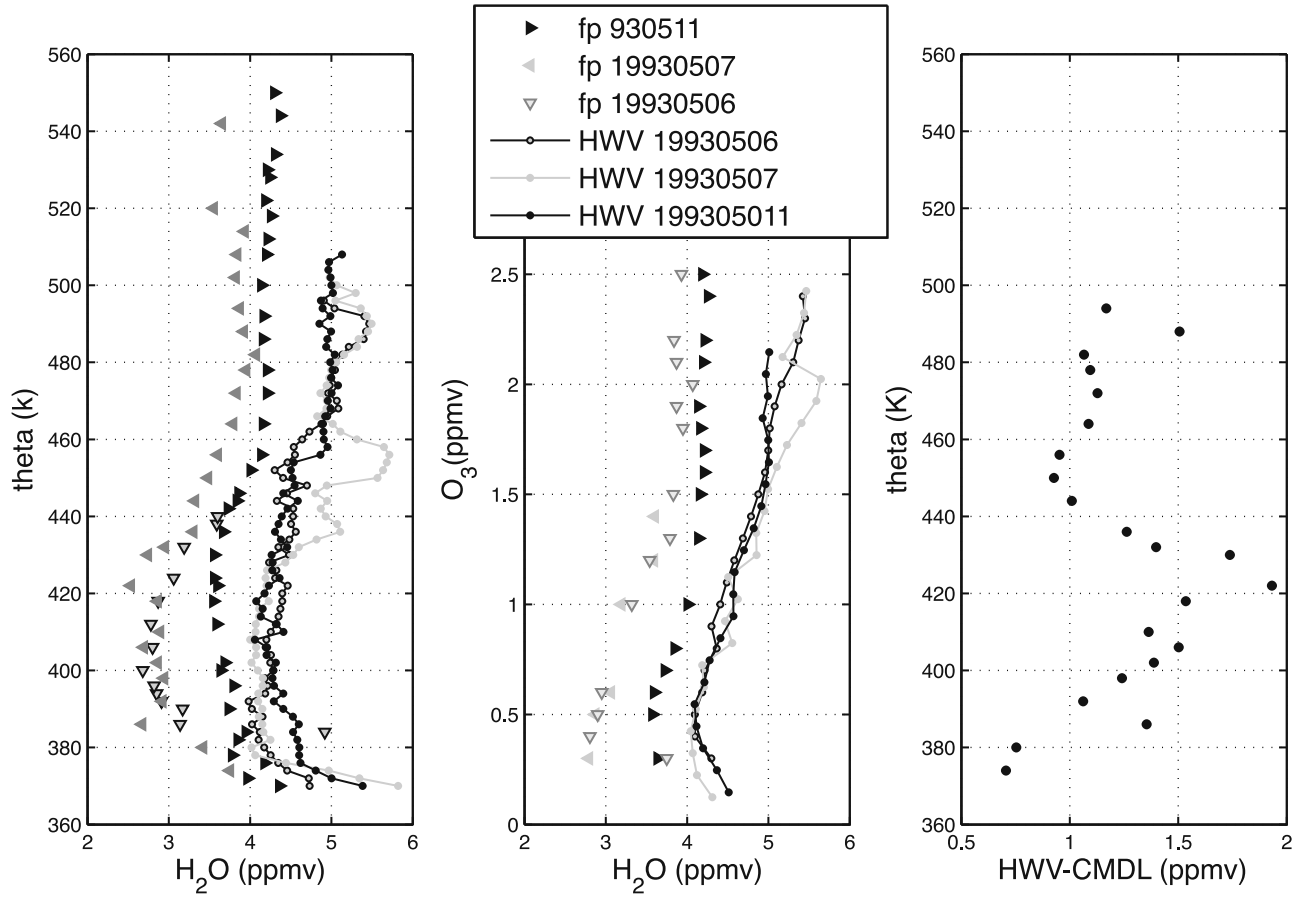

Figure 11. Intercomparison between HWV and the NOAA CMDL frost point hygrometer during the SPADE campaign. (left) CMDL and HWV water vapor profiles for the three flight dates indicated in the legend versus theta. (middle) Ozone−water vapor correlation. In this plot NOAA frost point ozone is used as the vertical coordinate for the CFH data and NOAA ozone [*Proffitt and McLaughlin*, 1983] is used as the vertical coordinate for the HWV data. (right) Difference between HWV and CMDL versus theta for flights on May 6, 1993.

19930331 and 19930404. The left panel shows the mean profiles and the right panel shows the difference in ppmv.

[38] These intercomparisons are consistent with those shown previously. However, HWV data were taken on average about ten days later than the CMDL data. In response to concerns that this intercomparison might be impacted by changes in stratospheric humidity over time, we looked at tropical profiles from HALOE taken in early March and early April. The HALOE data above about 420 K suggest a potential drying out of the tropical stratosphere by about 0.5–0.7 ppmv over the entire month of March. If this decrease in humidity were linear in time, it might increase the difference plotted in Figure 10 by about 0.2–0.3 ppmv, making the difference in slightly better agreement with recent comparisons between HWV on the WB57 and CFH.

3.7. SPADE Intercomparisons

[39] Shortly after the CEPEX campaign, HWV participated in the SPADE campaign from Moffett Field, CA (37.4°N, 122.1°W) during which CMDL was launched from Crows Landing, CA (37.4°N, 121°W). The results of three intercomparisons, shown in Figure 11, are in general agreement with the CEPEX intercomparisons. By using ozone as the vertical coordinate, we show that the variability exhibited in the three frost point profiles is most likely instrumental, with the CMDL profile on 11 May the outlier. Because of the sparseness of data on the 7th, and the anomalous profile on the 11th, we use only data from the 6th to determine the representative difference between HWV and CMDL during SPADE shown in Figure 11 (right).

3.8. STRAT and POLARIS Intercomparisons

[40] The results of the water vapor intercomparisons and the observed differences between HWV and HALOE during the Pre-AVE mission provide motivation to find additional intercomparison opportunities with HALOE data in the tropics. While tropical in situ data are sparse, the six tropical water vapor profiles on 19951105, 19960213, 19960801, 19960808, 19961211, and 19970923 provide potential additional opportunities for HALOE intercomparisons, depending on spatial and temporal overlap criteria. For the correlative measurements used in the HALOE validation paper [*Harries et al.*, 1996], overlap constraints "take pragmatic account of what coincidences are actually available." All of the intercomparisons shown by *Harries et al.* [1996] were at midlatitudes and northern latitudes, with

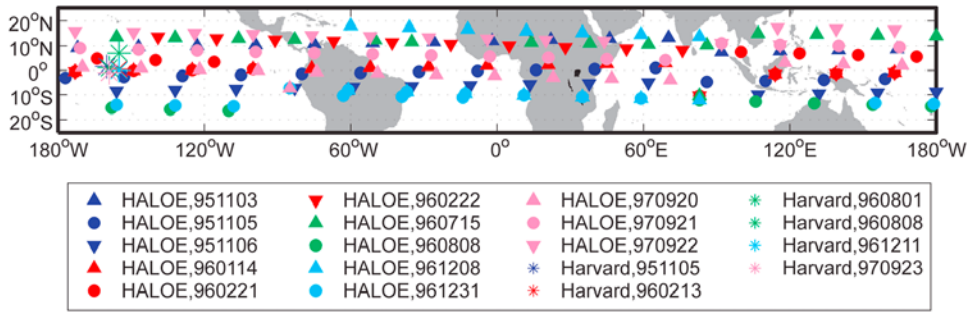

Figure 12. The geographic distribution of the in situ flight profiles and HALOE observations used in the intercomparison in the tropics. The points representing the locations of the HALOE profiles are color-coded to match the points plotted as asterisks representing locations of the in situ data to which they are compared.

typical constraints of 3° latitude, 15° longitude, and two days. For the tropical intercomparisons presented here, we assert that the homogeneity of water vapor mixing ratios in the lower tropical stratosphere allow these overlap constraints to be relaxed. Our strategy, therefore, is to bracket the in situ flight dates and location with a distribution of HALOE profiles to provide confidence that any observed differences between the in situ and HALOE profiles represent true measurement differences and not differences caused by temporal or geographic inhomogeneities in the atmosphere.

[41] We show in Figure 12 the locations of the in situ and HALOE profiles that we use in the intercomparison. The locations are color-coded to indicate the match between the in situ profiles and the HALOE data used for comparison. Dates for each of the profiles are given in the legend. We have clearly relaxed the geographical constraint for most of the intercomparisons, and in most cases the temporal

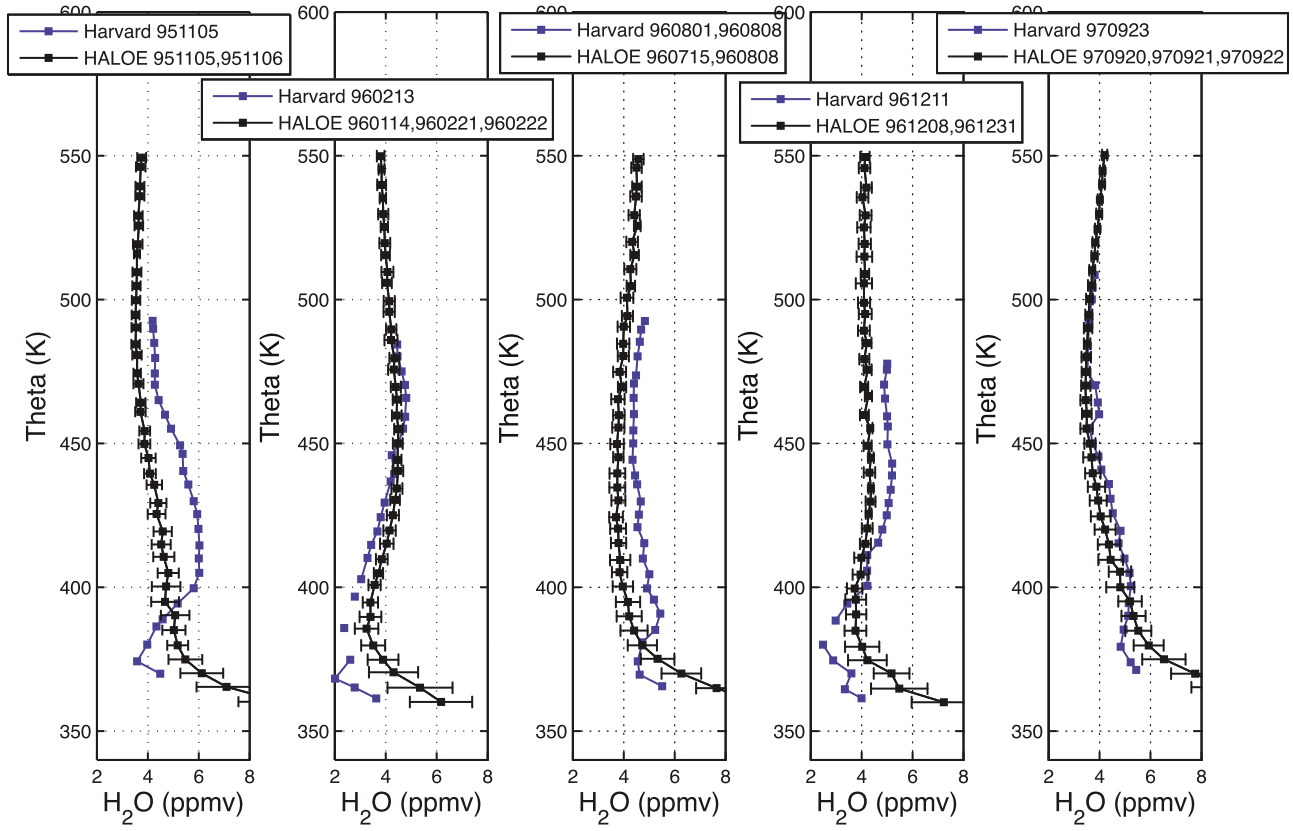

Figure 13. Intercomparison of Harvard and HALOE water vapor for tropical profiles taken during the STRAT and POLARIS campaigns from November 1995 to September 1997. For each of the in situ profiles plotted here, the corresponding HALOE profiles have been taken over two or three days and cover a significant fraction of the latitude circle. The HALOE mixing ratios have been corrected as in Figure 9.

constraint as well. Above about 400 K, we rely, based on our understanding of the dynamics in the tropical stratosphere, on the relative homogeneity therein. While the lower tropical stratosphere is isolated from midlatitudes as described by the "tropical pipe" model of the atmosphere [*Plumb*, 1996], weak isentropic mixing of midlatitude air into the tropical stratosphere does occur up to about 500 K [*Minschwaner et al.*, 1996; *Volk et al.*, 1996]. Nevertheless, the air that gets mixed into the lower tropical stratosphere does not have a significant impact on water vapor mixing ratios [*Weinstock et al.*, 2001]. The homogeneity in the tropical stratosphere is empirically confirmed by the self-consistency of the HALOE profiles used to compare with the respective in situ flight profiles and illustrated by the plotted 1-σ statistical standard deviations.

[42] For the five intercomparisons we include HALOE profiles from dates both before and after the in situ flight date when they are available. Starting with the intercomparison with the 951105 STRAT profile, we see in Figure 13 that the three days of HALOE profiles cover the entire tropics from 10°S to 10°N pretty well. For the intercomparison with 19960213 flight data, the HALOE profiles provide excellent coverage over the tropics from about 3°S to 13°N, with the dates bracketing the in situ flight date, but not very close in time. The biggest potential problem in this case would show up in the 390 to 450 K region where in over a month's time the signature of the seasonal cycle of water would propagate upward about 15 K. If this were to significantly affect the intercomparison, it would show up as a difference between the January and February HALOE profiles. We will see shortly that this is not the case. For the August intercomparison, we have to again rely on HALOE profiles two to three weeks prior to the in situ flights along with some that have exact temporal coincidence with the August 8 flight, although the latitude range only extends south to about 11°N. Nevertheless, the inclusion of the 15 July HALOE profiles, all in the southern hemisphere, again provides evidence of water vapor homogeneity in the tropics. For the December intercomparison, the HALOE flight dates bracket the in situ flight date but only provide latitudinal coverage below 10°S and above 12°N, thus providing among the intercomparisons the poorest spatial overlap with the in situ flight profile. For September 1997, the HALOE flights provide excellent tropical coverage and good temporal coincidence, comprising the three day period prior to the in situ flight date, thus providing an excellent intercomparison opportunity.

[43] To evaluate the results of these intercomparisons, we first focus on the tropopause region, (360–400 K) where the HALOE profiles, unlike the in situ profiles, exhibit similar gradients with potential temperatures in all seasons. It is important to address whether this inconsistency with the in situ data is instrumental or atmospheric. We start with the February profiles, for which the result of the intercomparison is very different than for the Pre-AVE intercomparison shown in Figure 9. The in situ profile, which is apparently impacted by local convective dehydration [*Weinstock et al.*, 2001], is lower than the HALOE profiles by almost a ppmv near 390 K, and remains about 0.5 ppmv lower up to 420 K, with the profiles converging above 440 K. In the late January 2004 Pre-AVE intercomparison, HALOE is lower by 1.5 to 0.5 ppmv from 390 to 420 K. This intercomparison strongly suggests that to intercompare with HALOE below about 410 K, direct planned intercomparisons are probably necessary. We are accordingly hesitant to attribute these observed gradient differences in the tropopause region to HALOE's inability to resolve the gradients observed by HWV.

[44] To evaluate differences between the HALOE and in situ profiles above the tropopause region, we average the HALOE data in 2 K potential temperature bins and plot in Figure 14 the percent and absolute difference from the in situ profiles with which they are being compared, color-coded by the dates of the in situ profiles. Starting from the tropopause region at about 370 K and extending to about 450 K each of the five intercomparisons exhibit self-consistent yet distinctive behavior. Around 400 to 420 K the HALOE profiles range from being about 30% low to 30% high, with the winter intercomparisons in February and December exhibiting different behavior from one another starting above about 400 K. We can attribute this to the very dry character of the in situ February flight in the tropopause region. However, even near 440 K a very large difference is maintained between the December and February intercomparisons.

[45] The most robust and important result of these intercomparisons is that above 460 K the spread in the difference data for all the intercomparisons starts decreasing and reaches a minimum of about 0–20%, with HALOE lower, at about 480 K. On average, HALOE is about 10% or 0.5 ppmv lower than the in situ value above ~460 K. This average absolute difference is in agreement with results from the Pre-AVE intercomparison but smaller than that shown in Figure 1 of the SPARC report. Accordingly, in this region, HALOE and HWV agree within their error bars and the disagreement is much less than that exhibited between HWV and CMDL. Nevertheless, especially below 460 K, this 20% is not randomly distributed among the five intercomparisons. The plotted differences vary somewhat systematically from one intercomparison to another. Unlike in the tropopause region, these variations do not appear to be atmospheric in origin.

[46] A comparison of these results with those from Pre-AVE suggests that in the tropopause region, atmospheric conditions local to the aircraft flight track can cause significant differences between the in situ and remote measurements, especially when overlap conditions are relaxed. Mixing reduces these inhomogeneities between about 390–420 K, allowing for more reliable intercomparisons. However, even above 420 K, there are systematic variations in the differences between HALOE and HWV possibly from atmospheric inhomogeneities or potential issues with the HALOE retrieval. The relatively steady measurement difference of about 0.5 ppmv above 460 K, however, is inconsistent with the general agreement between HALOE and CMDL [e.g., *Harries et al.*, 1996] and the ~1.5 ppmv disagreement between HWV and CMDL.

3.9. A Laboratory Intercomparison at the AIDA Chamber

[47] A formal water vapor intercomparison campaign, AquaVIT, was carried out at the AIDA aerosol chamber in Karlsruhe, Germany [*Kamm et al.*, 1999] in October 2007. The reported goal of this intercomparison was to

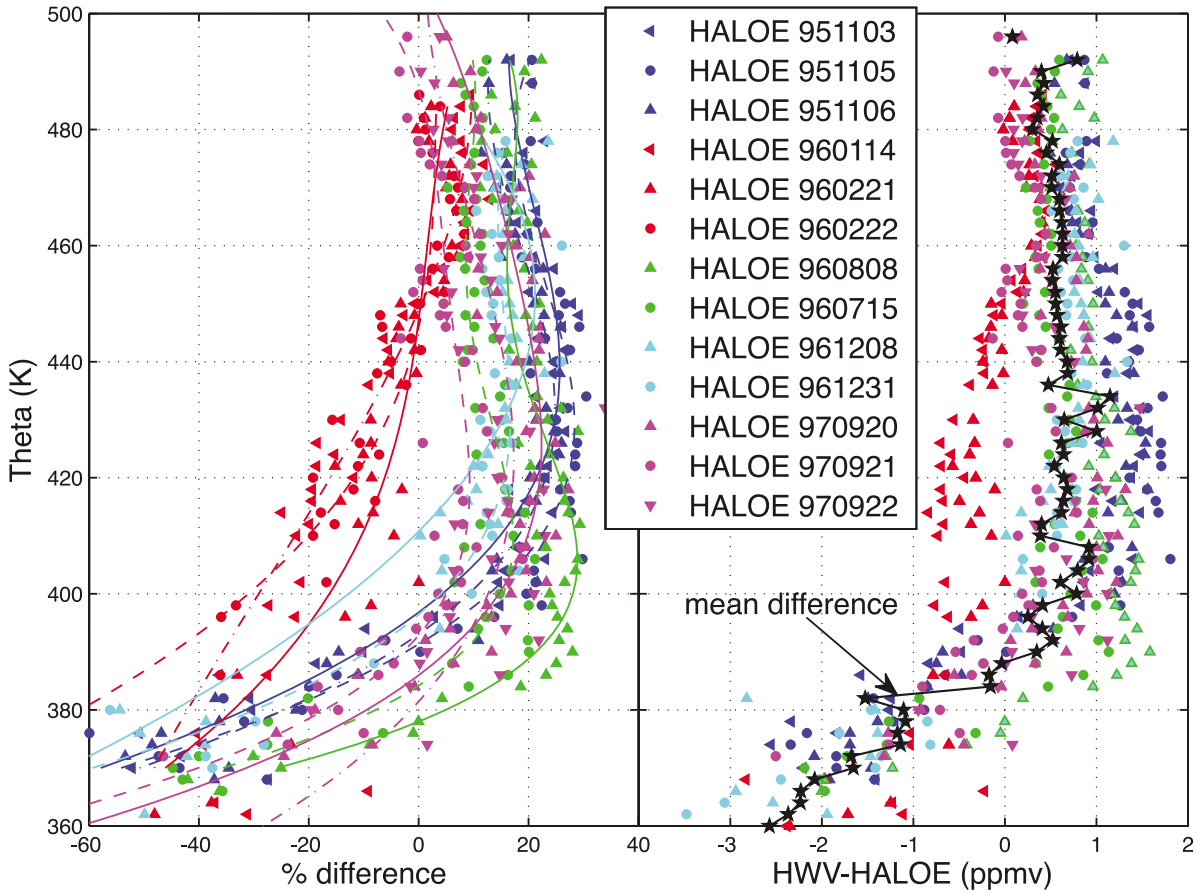

Figure 14. Plot of (left) percent difference ((in situ-HALOE)/in situ) and (right) absolute difference for the tropical profiles taken during the STRAT and POLARIS campaigns. The data are averaged in 2 K potential temperature bins and color-coded to differentiate the profiles taken during different seasons. Polynomial fits to the data are included in the left plot to help distinguish seasonal differences. An overall mean difference is provided in the right plot.

provide the opportunity to intercompare the performance of a number of instruments simultaneously measuring the humidity of air from a single homogeneous source over a wide range of pressures, temperatures, and water vapor mixing ratios (D. W. Fahey et al., Summary of the AquaVIT water vapor intercomparison, available at https://aquavit.icg.kfa-juelich.de/AquaVit/, 2009). Briefly, instruments were mounted both in and external to the chamber for sampling. In either case, significant attempts were made to provide the sampling conditions requested for each instrument, with the goal of eliminating sampling errors. For the data we present here, a typical experiment consisted of a constant temperature run in which pressure was stepped up at intervals of approximately one hour from a minimum pressure (50 to 100 hPa) to a maximum pressure (300–500 hPa) and back down, typically for a total of 7 intervals. Chamber temperatures were stepped down daily from 243 K for the first experiment to 185 K for the last, with the corresponding range of water mixing ratios decreasing as well. In order to ensure a blind intercomparison, a careful data analysis protocol was followed such that no instrument PI had access to data from other instruments until after an agreed upon date on which all data had been archived.

[48] The interpretation of in-flight instrument performance based on the results of a laboratory intercomparison can only be reasonably made after thoroughly evaluating the differences between the laboratory and flight operating conditions for each instrument. This is true for both instruments mounted inside the chamber and for those mounted outside the chamber that sampled chamber air through a heated sample tube. We address the correspondence between laboratory and flight conditions, specifically for HWV, with the goal of further validating the in-flight performance of HWV and assessing the systematic in-flight differences that have already been presented.

[49] Here we evaluate comparisons of data among a select group of instruments. They include the AIDA TDL, a facility instrument mounted inside the chamber [*Ebert et al.*, 2005; *Möhler et al.*, 2003]. The others, included because of the significant body of atmospheric data they have provided, are JLH, also mounted inside the chamber, CFH, and the Fast In situ Stratospheric Hygrometer (FISH) [*Zöger et al.*, 1999], both mounted external to the chamber. We include CFH and JLH because we participated in the comparison at the AIDA chamber in order to specifically address observed in-flight differences with those instruments. While we have not provided details of in-flight differences with JLH, their recent TC4 flight data agrees much better with CFH than with HWV at low water mixing ratios. We include a FISH instrument because that

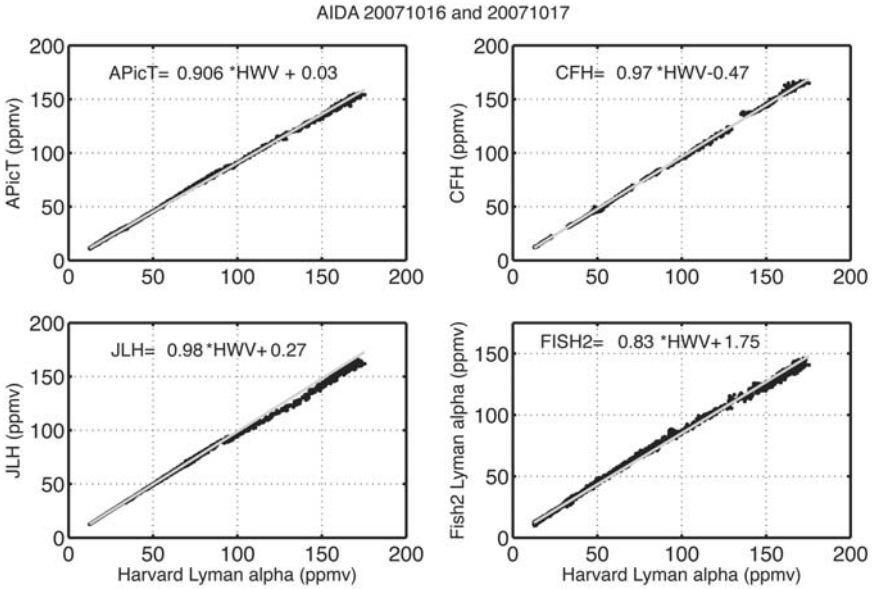

Figure 15. Intercomparison data for all the data submitted for the experiments in aerosol-free air in the AIDA chamber on 20071016 and 20071017. Dashed lines representing linear least squares fits to the data averaged over ten seconds are included.

instrument has provided a significant body of stratospheric data, has participated in many in-flight intercomparisons, and uses the same principle of detection as HWV, photofragment fluorescence excited by Lyman α. There were two FISH instruments participating in AquaVIT. We include data from FISH2, which, along with CFH, shared a heated sample tube with HWV.

[50] We have classified the sources of in-flight measurement differences as calibration-, sampling-, or instrument-artifact related. As already stated, for AquaVIT the sampling differences term should be negligible, especially at high humidities. We show in Figure 15 plots of HWV versus the AIDA TDL, JLH, CFH, and FISH2 using data from experimental runs with relatively high water vapor. This allows us to first compare how well these instruments agree in the laboratory versus in flight. It also allows us to identify instrument differences at low water vapor as being caused by calibration differences and/or by instrument artifacts. The plots show that while HWV agrees very well with the CFH and JLH instruments over a wide range of mixing ratios, it differs by about 10% with the AIDA TDL and more than 15% with FISH2. We interpret these as calibration differences that can best be evaluated by direct laboratory calibrations and intercomparisons without the AIDA chamber. Also, when making atmospheric measurements as total water instruments, both the HWV and FISH detection modules are operating in flight under the same pressure and temperature conditions as they do in the laboratory. Accordingly, direct laboratory intercomparisons between these two instruments could resolve the differences observed during AquaVIT, and by extension in flight. Since the mission, these differences have in part been resolved by laboratory calibration of the FISH2 instrument at high pressures (C. Schiller, personal communication, 2009). We can now use the slopes from the least squares fits to normalize the data from each instrument to HWV. Measurement differences at low water using normalized data provide the best indicator of potential instrumental artifacts at low water vapor, since we have effectively removed the contribution due to differences in instrument calibration.

[51] For HWV, however, the low water measurements may also be impacted by inadequate flow or small leaks in our instrument duct. During the week prior to the formal intercomparison experiments, diagnostic tests showed conclusively that the maximum flow rates attainable through HWV and the inlet tube connecting it to the AIDA chamber, especially at pressures of 100 hPa and below, were insufficient to eliminate positive systematic offsets from outgassing surfaces and small leaks at the lowest water vapor mixing ratios. Because the capacity of the pumps that draw air from the AIDA chamber through our instrument have a maximum volume displacement, lower mass flow rates occur at lower pressures and the potential for contamination is therefore greater under these conditions. A flow model was developed to estimate a flow- and pressure-dependent offset that contributed to the measured water vapor. Data were archived only if this correction was less than 20% of the measured value.

[52] We utilize the calibration relations presented in Figure 15 to normalize the low water measurements from experiments on 20071018 and 20071019, even though those water measurements were low enough that calibration errors only have a minor impact on the results. We plot in Figure 16 differences between Harvard and the other instruments for the experiments on these two dates. Note that there was no temporal overlap between our archived data and JLH data on 20071019, and only limited overlap on 20071018. We include in Figure 16 a plot of the modeled offset correction as an indicator of the magnitude of the potential systematic error caused by insufficient mass flow.

[53] How well do the data plotted in Figures 15 and 16 constrain a possible positive offset in HWV and to what extent do the observed differences account for the differences observed in flight? While there is reasonably good

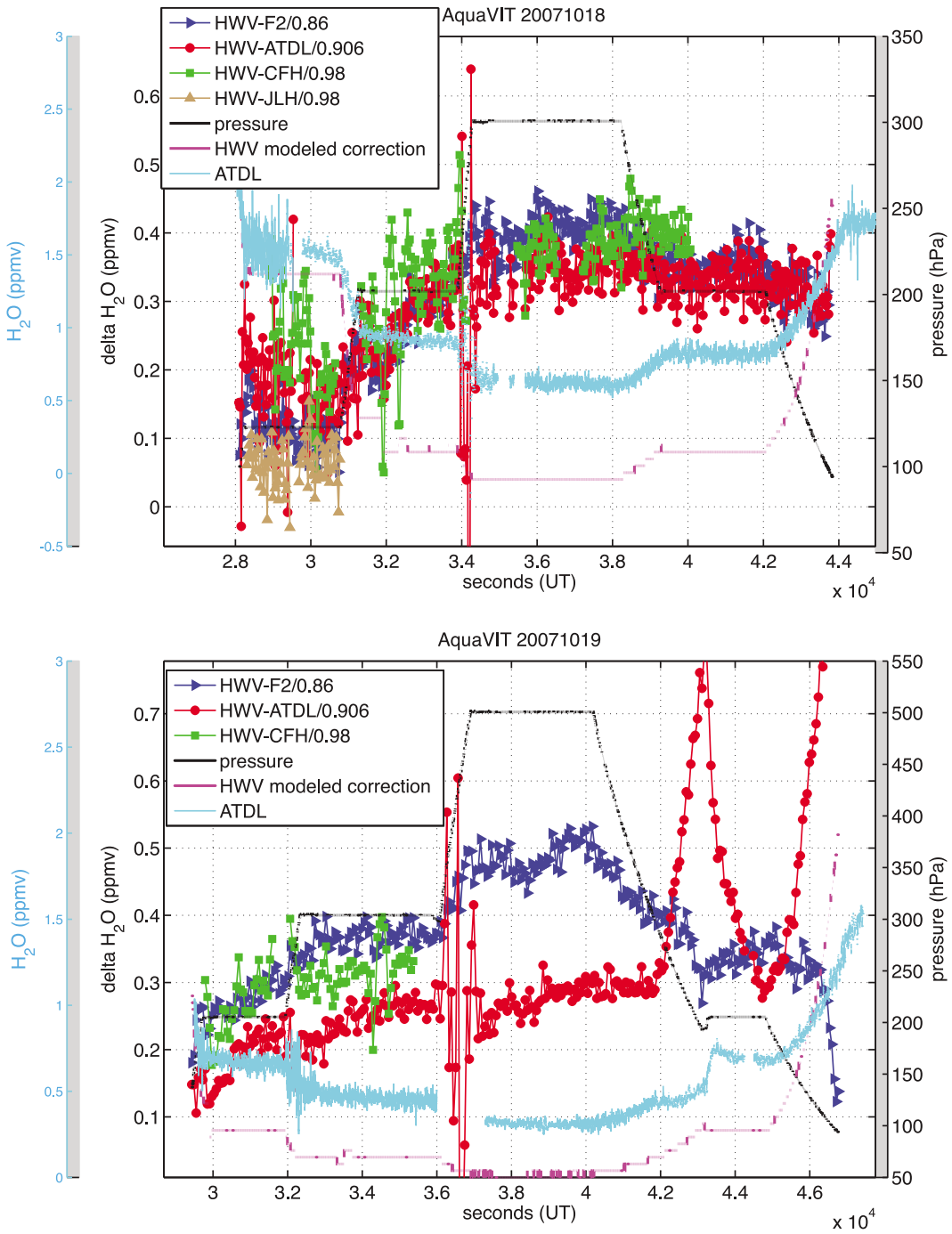

Figure 16. Plot of measurement differences versus time between Harvard water vapor and (top) the four other instruments of particular interest, FISH2, the AIDA TDL, CFH, and JLH on 18 October and (bottom) three of those on 19 October. There were no overlapping data for HVD and JLH on 19 October. On 19 October, the chamber wall temperature was maintained near 185 K, thus providing a frost layer on the wall and control of the chamber humidity. With each pressure change (black line), a steady wall temperature was typically reached after 10–15 min. For reference, we also plot both water vapor measured by the AIDA TDL and the correction to HWV from the flow model.

consistency throughout the three days of experiments, there is some variability in the observed differences, likely due to unidentified sensitivities to the sampling conditions, including flow, pressure, temperature or time, of one or more of the instruments. Additionally, over or under correction of HWV by the flow model could contribute to some of the observed variability.

[54] We start by looking at data exhibiting the smallest differences. For the low-pressure segment of data taken on 20071018, the smaller observed differences could result from an overestimation of the low flow correction factor.

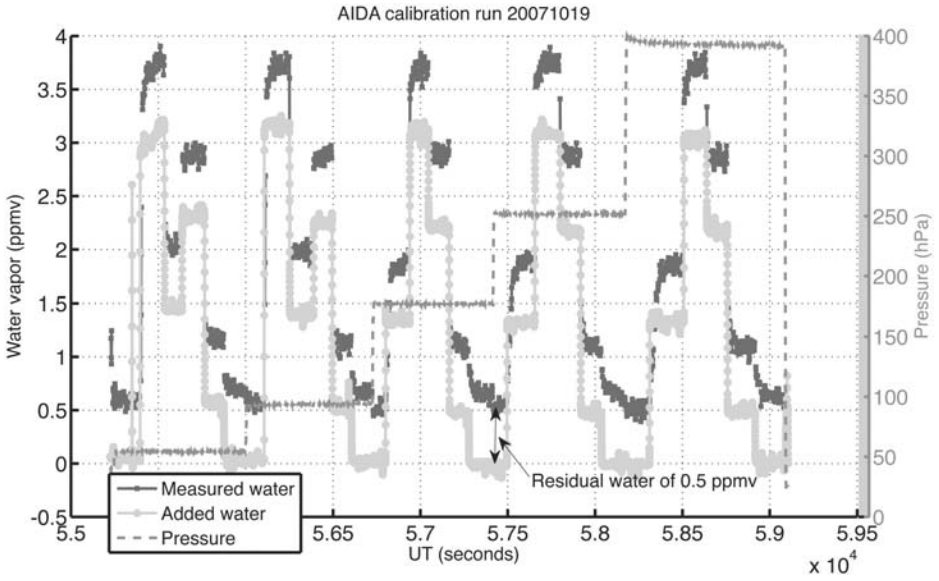

Figure 17. Calibration run in air during the AIDA campaign, covering the identical low water region sampled during the run that day over a wide range of pressures. Residual water vapor of 0.5 ppmv is observed.

However, even if the correction were reduced by 0.2 ppmv, thus bringing the plotted differences in that segment up to those in the next data segment, the difference between HWV and JLH would be about 0.25 ppmv. For the lowest pressure data on the 20071019, taken at 200 hPa, which also show similar small differences, the correction to HWV is less than 0.1 ppmv. Since this pressure regime is close to the atmospheric pressure regime of interest, these results are important.

[55] Typically, with calibration differences accounted for, HWV measures about 0.30 ± 0.05 ppmv higher than the AIDA TDL. The difference from FISH2 is slightly higher, about 0.40 ± 0.10 ppmv. The higher scatter in the measured low water difference with FISH2 is probably caused by variability of its calibration along with HWV's. While the difference between the HWV and JLH data is more variable, it is typically 0.05–0.10 ppmv less than that with the AIDA TDL. Based on limited intercomparison data on the 18th and 19th, HVW is about 0.35 ± 0.05 ppmv higher than CFH. Most importantly, the conclusion we can draw is that the respective differences during low water measurements at AIDA between HWV and both CFH and JLH are nowhere near the in-flight differences historically observed with the CFH instrument and with JLH during TC4.

[56] Immediately following the last experiment on the 19th, (as well as at other times during the week), a low water calibration run of HWV was carried out over a range of pressures, the results of which are shown in Figure 17. Measured water vapor is plotted in black, and added water vapor, determined from the vapor pressure of water in a bubbler, gas flow rates, and pressure measurements [*Weinstock et al.*, 2006a, 2006b], is plotted in gray. The difference between the two curves is about 0.5 ppmv and represents, as in Figures 2 and 3, the sum of water vapor in the carrier gas, water vapor outgassing from surfaces in the chamber, inlet tubes, and instrument, as well as any an instrument offset. While there were no other independent measurements of the humidity of this air, these measurements illustrate that any offset must be less than 0.5 ppmv. Realistically, manifolds of dry air have water vapor mixing ratios of at least 0.2 ppmv.

[57] The AquaVIT results are consistent with the low-water calibrations performed at Harvard that constrained any potential Lyman-α instrument artifact to $\sim$<0.2 ppmv in air. Furthermore, we have shown that these laboratory results are applicable to flight data. Therefore, we conclude that the differences observed in flight between HWV and CFH over the past two decades and most recently with JLH [see *Gensch et al.*, 2008] are not explained by the AquaVIT results.

4. Implications of Water Vapor Measurement Discrepancies

[58] As we have previously noted, accurate water vapor measurements in the UT/LS are needed to constrain the seasonal cycle of entry-level stratospheric water vapor. Satellite data, with global coverage over all seasons, whether from the HALOE or MLS instruments, provide extensive tropical data for testing models of the mechanisms that control stratospheric water vapor. As shown in Figure 9, careful intercomparisons between HWV, CFH and HALOE were carried out in the lower tropical stratosphere during the Pre-AVE campaign. Similar to *Weinstock et al.* [2001], we determine the stratospheric age of a sampled air parcel using CO measured on the WB-57 and a simple photochemical model, and use that age to identify the date the air crossed the 390 K surface, a surrogate for the tropical tropopause. As before, the validity of the calculated stratospheric age of the air sampled during Pre-AVE is tested by comparing it, as shown in Figure 18a with the stratospheric age of the air from boundary condition or entry-level CO_2, $(CO_2)_{el}$. $(CO_2)_{el}$ is derived from the average of the Mauna Loa and Samoa surface stations, operated by the Global Monitoring Division

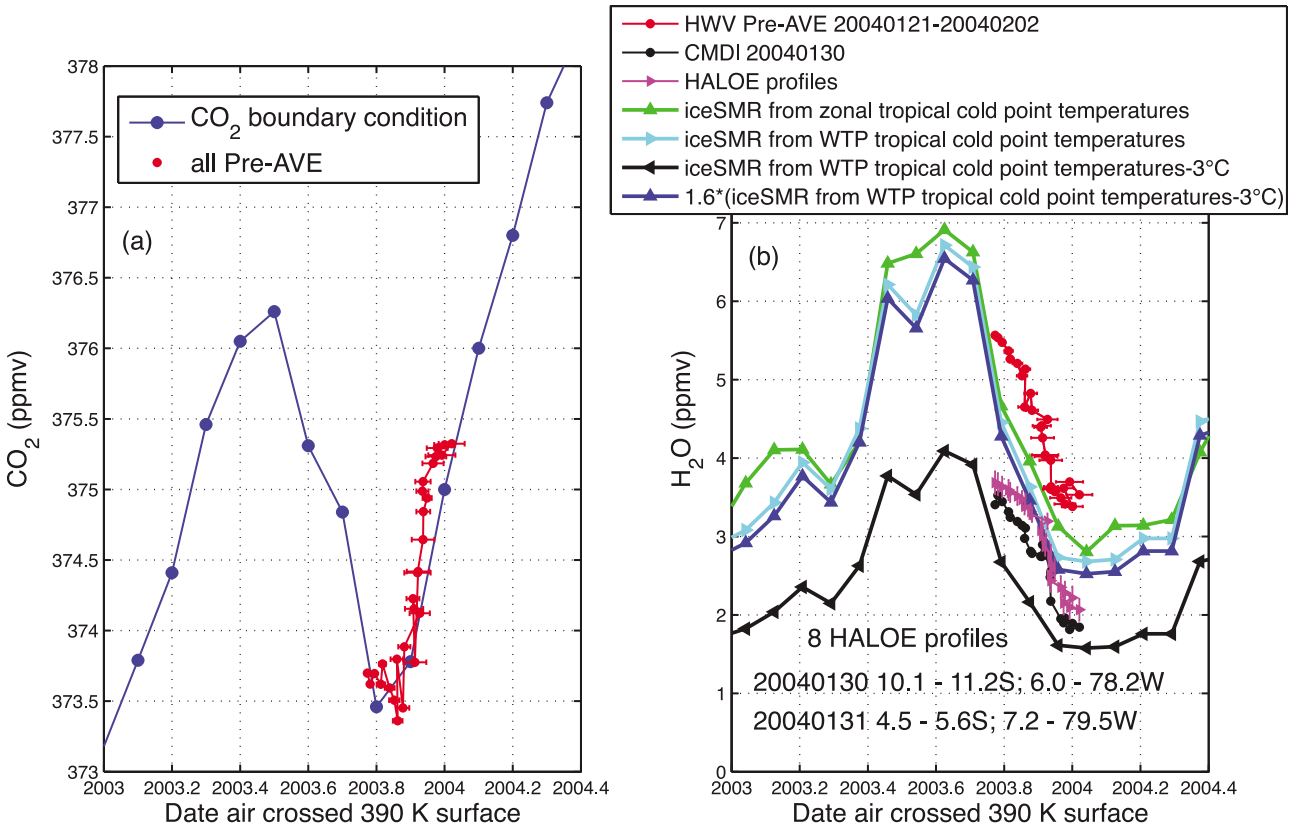

Figure 18. Mixing ratios measured above the tropical tropopause plotted versus the date that the sampled air mass crossed the 390 K isentrope as determined by measured CO and the photochemical model: (a) carbon dioxide and (b) water vapor. Included in Figure 18a are $(CO_2)_{el}$ boundary condition values determined from tropopause measurements as described in the text. Included in Figure 18b are profiles of CMDL data plotted in Figure 10 as well as the corrected HALOE profiles binned and averaged identically to the Harvard in situ data. For comparison to represent $(H_2O)_{el}$, we plot ice saturation mixing ratios derived from monthly averaged radiosonde cold-point temperatures in the tropics, along with those derived from cold-point temperatures from 22 radiosonde stations that are distributed throughout the tropics, and those derived from data from nine stations in the western tropical Pacific (WTP). Adjustments to the data from the WTP as described in the legend are provided to enable comparison with the work of *Fueglistaler et al.* [2005].

at the NOAA Earth System Research Laboratory, with an assumption of a 60 day transit time from the surface to 390 K [*Boering et al.*, 1994]. In the analysis presented by *Weinstock et al.* [2001], the isentropic mixing of midlatitude air into the lower tropical stratospheric had to be included in order to match the CO_2 derived from in situ data with that from surface measurements. However, including that mixing only marginally impacted the entry-level water vapor mixing ratios derived from the model. Accordingly, we do not include midlatitude mixing in the model used to derive the Pre-AVE data plotted in Figure 18a. The excellent agreement between entry-level CO_2 derived from surface CO_2 and that derived from in situ measurements in the lower tropical stratosphere in combination with the photochemical model validates the model for use in determining the seasonal phasing of $(H_2O)_{el}$ when isentropic mixing from midlatitudes is minimal. In Figure 18b we provide context for the $(H_2O)_{el}$ derived from water vapor measured in the lower tropical stratosphere during Pre-AVE, using monthly averaged ice saturation mixing ratios derived from two different sets of radiosonde cold-point temperatures taken between 15°S and 15°N, which provide an alternative representation of $(H_2O)_{el}$. One set of stations nominally provide a mixing ratio assuming zonally averaged ascent, while the other, investigates the possibility of preferential ascent and dehydration in the western tropical Pacific.

[59] We provide further context by also plotting ice saturation mixing ratios derived from temperatures 3 degrees below the western tropical Pacific cold point temperatures. This temperature drop represents the average temperature difference between the Eulerian cold point temperatures and those corresponding to the minimum ice saturation mixing ratios experienced during the Lagrangian back trajectories from *Fueglistaler et al.* [2005]. Additionally, we include the possibility that, consistent with both in situ water vapor data in clear air and clouds in the TTL, dehydration events do not reduce water down to ice saturation. Taking this into account can potentially bring the $(H_2O)_{el}$ from the Fueglistaler model into better agreement with that derived from HWV data.

[60] By including the different representations of the water vapor seasonal cycle derived from cold-point

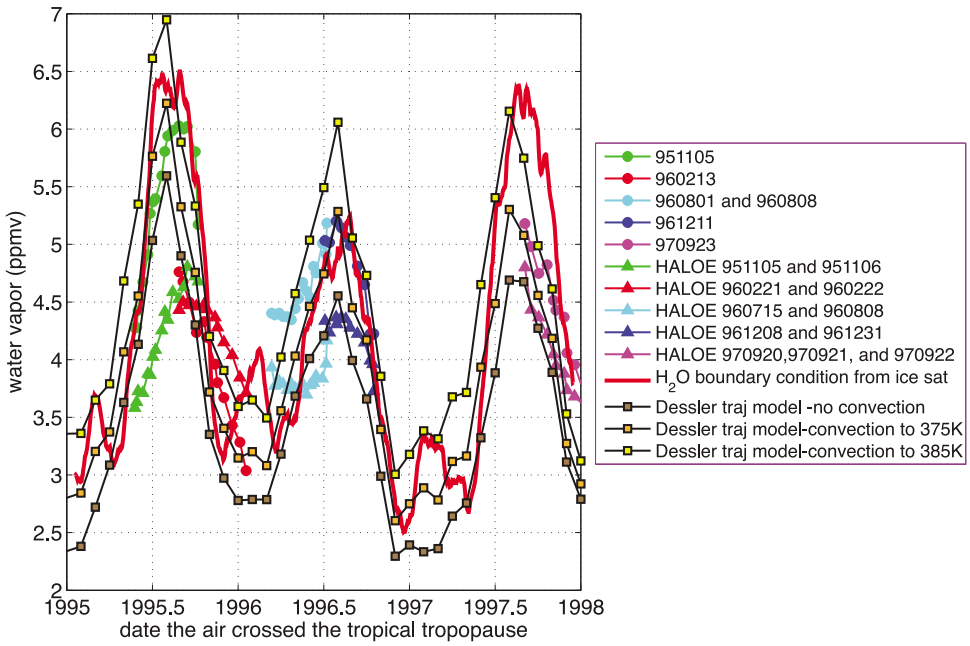

Figure 19. Comparison of the seasonal cycle of water vapor entering the lower tropical stratosphere derived from HWV and HALOE data and calculated from ice saturation at mean daily zonal tropical cold-point temperatures and from a Lagrangian trajectory model adapted by *Dessler et al.* [2007] to include convection.

tropopause temperatures we can better appreciate the difference between using Harvard in situ data as opposed to frost point or HALOE data when evaluating dehydration mechanisms. The HALOE data, 1.0–1.5 ppmv lower than the Harvard data, or that of frost point data, 1.5–2 ppmv lower than the Harvard data, are clearly more consistent with the Fueglistaler model. Agreement with the Harvard data would clearly require a mechanism that either provides significantly less dehydration or meaningful convective ice evaporation, which is not suggested by the Fueglistaler model.

[61] In spite of the reasonable agreement between $(H_2O)_{el}$ derived from in situ water vapor measurements and $(H_2O)_{el}$ from radiosonde data, and consistent with a conclusion stated by *Dessler* [1998], we do not suggest that it supports a specific dehydration mechanism. In fact, even with the temperatures averaged over the stations heavily weighted toward the western tropical Pacific, the Eulerian model cannot properly take into account the actual geographical distribution of air mass trajectory entry points. Nevertheless, it does clearly illustrate how evaluation of the viability of a specific dehydration mechanism depends on the water vapor data selected for comparison with the model.

[62] We now revisit the seasonal cycle of water derived from tropical profiles measured by HWV on the ER-2 during the STRAT and POLARIS campaigns as described by *Weinstock et al.* [2001]. We duplicate in Figure 19 the right plot of Figure 9 of *Weinstock et al.* [2001]. We include as well the boundary condition water vapor derived from binned and averaged HALOE water vapor data, also using the relationship between theta and stratospheric age derived from the photochemical model. Additionally, we include data from the Lagrangian trajectory model in which respectively, no convection, convection to 375 K, and convection to 385K are included [*Dessler et al.*, 2007]. These results show that when no convection is included, the model output agrees more with HALOE data but when convection is included to about 375K or 385 K, it agrees more with the in situ data, depending on the year. So here we see that the Lagrangian model matches HALOE data without convection but requires convection to 375 or 385 K to match the seasonal cycle derived from in situ data. These results are complementary to those presented in Figure 18 in which the results from the Fueglistaler model agree with HALOE data during Pre-AVE but would better match the in situ data if the air parcels on the back trajectories in the model were only dehydrated to a relative humidity over ice consistent with HWV supersaturation data.

5. Conclusions

[63] As can be seen from Figures 18 and 19, a critical appraisal of the accuracy of the instruments that provide the preponderance of stratospheric water vapor data is necessary if they are to be used to monitor changes in stratospheric humidity and to help evaluate the mechanisms that control it. The intercomparisons in Figures 8–11 that illustrate the large difference between the Harvard hygrometer and the CMDL and CFH frost point hygrometers is consistent with other intercomparisons and with the results summarized in Figure 1 of *Kley et al.* [2000]. In this manuscript we have provided details of the process of validating the in-flight accuracy of the HWV instrument. We believe that the described methodology, consisting of: (1) laboratory calibrations under flight conditions; (2) in-flight diagnostics that constrain potential systematic errors; and (3) in-flight intercomparisons that constrain potential systematic errors not diagnostically constrained, should serve as a model for

instrument validation. The significantly better agreement exhibited between HWV and CFH during the AQUAVIT campaign than in flight points to in-flight sampling errors and/or artifacts as the cause of the 1–1.5 ppmv systematic difference observed in flight. In this paper we have detailed both the validation methodology and diagnostic process that constrain HWV in-flight sampling errors or instrument artifacts to 0.25 ppmv or less.

[64] The intercomparison data presented throughout this manuscript illustrate a systematic difference between two groups of measurements. The frost point hygrometer generally exhibits better than 10% agreement with HALOE and MLS, and has often been used to validate those instruments. The instruments on the WB-57 have been shown in Figure 5 and 6 to be self-consistent and are significantly different from MLS and frost point in the tropopause region and lower stratosphere, and HALOE in the tropopause region. Even when the recommended 20% correction is made to the HALOE data in the tropopause region, differences up to 30% remain between 380 and 400 K. The systematic variability exhibited in the intercomparisons represented in Figure 13 is not consistent with systematic or random changes in the HWV detection sensitivity over time. Rather, it is suggestive of errors, most likely systematic, in analyzed HALOE data. On the other hand, the differences do decrease with decreasing pressure, to about 10% or ~0.5 ppmv, a significantly smaller difference than observed between HWV and CFH. Nevertheless, models used to identify and/or evaluate the mechanisms that control stratospheric water vapor often use HALOE data in the tropical tropopause region for validation. Accordingly, if the spread in the intercomparison differences in the tropopause region is indicative of a HALOE instrumental or analysis artifact, it must be understood.

[65] On the one hand, if the conclusion of this paper is accepted, i.e., that the in-flight validation of HWV has been demonstrated, there are two interpretations to the comparison of the modeled and measured values of $(H_2O)_{el}$. Either convection in the tropics reaching 375–385 K must be included in the trajectory model or the dehydration process on average leads to relative humidities over ice well above ice saturation. On the other hand, if CFH or MLS data are used, the conclusions are very different. Continued flight intercomparisons or follow-on multi-instrument laboratory efforts like AquaVIT can only be productive following a series of carefully orchestrated laboratory intercomparisons, each of which are designed to test the performance of instruments under simulated flight conditions. These intercomparisons can most efficiently be carried out two instruments at a time, with the core AquaVIT intercomparison instruments as prime candidates. Agreement in the laboratory must precede repeated flight intercomparisons.

[66] At the same time, advances in infrared laser technology and digital signal processing capabilities suggest a parallel approach should be followed in which an instrument is designed, developed, tested and calibrated in the laboratory under flight conditions, and satisfies all the criteria necessary to measure water vapor with sufficient precision and accuracy. While the resolution of current instrument differences are on the order of a ppmv, attaining accurate water vapor measurements of 1–2 ppmv in the TTL requires measurement uncertainties of no more than 0.1 ppmv. To pass this stringent requirement, an instrument needs to have better than 0.025 ppmv precision and accuracy. The final test of this instrument would consist of a flight intercomparison of two of them, thereby providing the opportunity to vary and compare sampling conditions and detection conditions in a systematic way.

[67] **Acknowledgments.** We greatly appreciate the efforts of the pilots and crews of the NASA ER-2 and WB-57 aircraft in making the missions we participated in so successful. We thank the AIDA support team and especially hosts Ottmar Möhler and Harald Saathoff for their hospitality and organizational efforts, and David Fahey and Ru-Shan Gao for serving as referees along with Ottmar Möhler during AquaVIT. We also thank the PIs for data we use in the intercomparisons: Robert Herman (JLH), Holger Vömel (CFH), Cornelius Schiller (FISH), and Volker Ebert (AIDA TDL). Thanks also go to the MLS and HALOE PIs and instrument teams for all the hard work and perseverance over many years and for providing easy access to their data. Thanks go to Holger Vömel for access to flight frostpoint data and helpful comments on the manuscript. Thanks also go to Andrew Dessler for supplying his model data used in Figure 19. Helpful comments from Stephan Fueglistaler, Eric Jensen, Leonard Pfister, and Ellis Remsberg and the three manuscript referees are gratefully acknowledged. Elliot Weinstock is grateful for the encouragement of David Fahey regarding the writing of this manuscript as well as his and Rushan Gao's challenging questions about HWV. The Harvard data in this manuscript would not exist without continuous support from the NASA Upper Atmospheric Research Program, most recently NASA grant NNG05GJ81G.

References

Boering, K. A., B. C. Daube Jr., S. C. Wofsy, M. Loewenstein, J. R. Podolske, and E. R. Keim (1994), Tracer-tracer relationships and lower stratospheric dynamics: CO_2 and N_2O correlations during SPADE, *Geophys. Res. Lett.*, *21*, 2567–2570, doi:10.1029/94GL01985.

Brewer, A. W. (1949), Evidence for a world circulation provided by the measurements of helium and water vapor distribution in the stratosphere, *Q. J. R. Meteorol. Soc.*, *75*, 351–363, doi:10.1002/qj.49707532603.

Dessler, A. (1998), A reexamination of the "stratospheric fountain hypothesis," *Geophys. Res. Lett.*, *25*, 4165–4168, doi:10.1029/1998GL900120.

Dessler, A. E., T. F. Hanisco, and S. Fueglistaler (2007), Effects of convective ice lofting on H_2O and HDO in the tropical tropopause layer, *J. Geophys. Res.*, *112*, D18309, doi:10.1029/2007JD008609.

Ebert, V., H. Teichert, C. Giesemann, H. Saathoff, and U. Schurath (2005), Fiber-coupled in situ-laser absorption spectrometer for the selective detection of water vapor traces down to the ppb-level, *Tech. Mess.*, *72*(1), 23–30.

Folkins, I., and R. V. Martin (2005), The vertical structure of tropical convection, and its impact on the budgets of water vapor and ozone, *J. Atmos. Sci.*, *62*, 1560–1573, doi:10.1175/JAS3407.1.

Fueglistaler, S., M. Bonazzola, P. H. Haynes, and T. Peter (2005), Stratospheric water vapor predicted from the Lagrangian temperature history of air entering the stratosphere in the tropics, *J. Geophys. Res.*, *110*, D08107, doi:10.1029/2004JD005516.

Gensch, I. V., et al. (2008), Supersaturations, microphysics and nitric acid partitioning in a cold cirrus cloud observed during CR-AVE 2006: An observation-modeling intercomparison study, *Environ. Res. Lett.*, *3*, 035003, doi:10.1088/1748-9326/3/3/035003.

Harries, J. E., J. M. Russell III, A. F. Tuck, L. L. Gordley, P. Purcell, K. Stone, R. M. Bevilacqua, M. Gunson, G. Nedoluha, and W. A. Traub (1996), Validation measurements of water vapour from the Halogen Occultation Experiment (HALOE), *J. Geophys. Res.*, *101*, 10,205–10,216, doi:10.1029/95JD02933.

Hintsa, E. J., E. M. Weinstock, J. G. Anderson, R. D. May, and D. F. Hurst (1999), On the accuracy of in situ water vapor measurements in the troposphere and lower stratosphere with the Harvard Lyman-α hygrometer, *J. Geophys. Res.*, *104*, 8183–8189, doi:10.1029/1998JD100110.

Holton, J. R., and A. Gettelman (2001), Horizontal transport and the dehydration of the stratosphere, *Geophys. Res. Lett.*, *28*, 2799–2802, doi:10.1029/2001GL013148.

Jensen, E., et al. (2005), Ice supersaturations exceeding 100% at the cold tropical tropopause: Implications for cirrus formation and dehydration, *Atmos. Chem. Phys.*, *5*, 851–862.

Jensen, E., et al. (2007), Formation of large (100 μm) ice crystals near the tropical tropopause, *Atmos. Chem. Phys. Disc.*, *7*, 6293–6327.

Kamm, S., O. Möhler, K.-H. Naumann, H. Saathoff, and U. Schurath (1999), The heterogeneous reaction of ozone with soot aerosol, *Atmos. Environ.*, *33*, 4651–4661, doi:10.1016/S1352-2310(99)00235-6.

Kley, D., J. M. Russell III, and C. Phillips (Eds.) (2000), *SPARC Assessment of Upper Tropospheric and Stratospheric Water Vapour*, World Clim. Res. Prog., Geneva, Switzerland.

Kuang, Z., G. C. Toon, P. O. Wennberg, and Y. L. Yung (2003), Measured HDO/H_2O ratios across the tropical tropopause, *Geophys. Res. Lett.*, *30*(7), 1372, doi:10.1029/2003GL017023.

Minschwaner, K., A. E. Dessler, J. W. Elkins, C. M. Volk, D. W. Fahey, M. Loewenstein, J. R. Podolske, A. E. Roche, and K. R. Chan (1996), The bulk properties of isentropic mixing into the tropics in the lower stratosphere, *J. Geophys. Res.*, *101*, 9433–9439, doi:10.1029/96JD00335.

Möhler, O., et al. (2003), Experimental investigation of homogeneous freezing of sulphuric acid particles in the aerosol chamber AIDA, *Atmos. Chem. Phys.*, *3*, 211–223.

Mote, P. W., K. H. Rosenlof, M. E. Mclntyre, E. S. Carr, J. C. Gille, J. R. Holton, J. S. Kinnersley, H. C. Pumphrey, J. M. Russell III, and J. W. Waters (1996), An atmospheric tape recorder: The imprint of tropical tropopause temperatures on stratospheric water vapor, *J. Geophys. Res.*, *101*, 3989–4006, doi:10.1029/95JD03422.

Newell, R. E., and S. Gould-Stewart (1981), A stratospheric fountain?, *J. Atmos. Sci.*, *38*, 2789–2796, doi:10.1175/1520-0469(1981)038<2789:ASF>2.0.CO;2.

Plumb, R. A. (1996), A "tropical pipe" model of stratospheric transport, *J. Geophys. Res.*, *101*, 3957–3972, doi:10.1029/95JD03002.

Proffitt, M. H., and R. J. McLaughlin (1983), Fast-response dual-beam UV absorption ozone photometer suitable for use on stratospheric balloons, *Rev. Sci. Instrum.*, *54*, 1719, doi:10.1063/1.1137316.

Randel, W. J., F. Wu, S. J. Oltmans, K. Rosenlof, and G. Nedoluha (2004), Interannual changes of stratospheric water vapor and correlations with tropical tropopause temperatures, *J. Atmos. Sci.*, *61*, 2133–2148, doi:10.1175/1520-0469(2004)061<2133:ICOSWV>2.0.CO;2.

Read, W. G., et al. (2007), Aura microwave limb sounder upper tropospheric and lower stratospheric H_2O and relative humidity with respect to ice validation, *J. Geophys. Res.*, *112*, D24S35, doi:10.1029/2007JD008752.

Read, W. G., et al. (2008), The roles of convection, extratropical mixing, and in-situ freeze-drying in the Tropical Tropopause Layer, *Atmos. Chem. Phys.*, *8*, 6051–6067.

Sayres, D. S., et al. (2009), A new cavity based absorption instrument for detection of water isotopologues in the upper troposphere and lower stratosphere, *Rev. Sci. Instrum.*, *80*, 044102, doi:10.1063/1.3117349.

St. Clair, J. M., et al. (2008), A new photolysis laser-induced fluorescence instrument for the detection of H_2O and HDO in the lower stratosphere, *Rev. Sci. Instrum.*, *79*, 064101, doi:10.1063/1.2940221.

Volk, C. M., et al. (1996), Quantifying transport between the tropical and mid-latitude lower stratosphere, *Science*, *272*, 1763–1768, doi:10.1126/science.272.5269.1763.

Vömel, H., et al. (2007a), Validation of Aura Microwave Limb Sounder water vapor by balloon-borne Cryogenic Frost Point Hygrometer measurements, *J. Geophys. Res.*, *112*, D24S37, doi:10.1029/2007JD008698.

Vömel, H., D. E. David, and K. Smith (2007b), Accuracy of tropospheric and stratospheric water vapor measurements by the cryogenic frost point hygrometer: Instrumental details and observations, *J. Geophys. Res.*, *112*, D08305, doi:10.1029/2006JD007224.

Weinstock, E. M., E. J. Hintsa, A. E. Dessler, J. F. Oliver, N. L. Hazen, J. N. Demusz, N. T. Allen, L. B. Lapson, and J. G. Anderson (1994), A new fast response photofragment fluorescence hygrometer for use on the ER-2 and Perseus remotely piloted aircraft, *Rev. Sci. Instrum.*, *65*, 3544, doi:10.1063/1.1144536.

Weinstock, E. M., et al. (2001), Constraints on the seasonal cycle of stratospheric water vapor using in situ measurements from the ER-2 and a CO photochemical clock, *J. Geophys. Res.*, *106*, 22,707–22,724, doi:10.1029/2000JD000047.

Weinstock, E. M., et al. (2006a), Measurements of the total water content of cirrus clouds. Part I: Instrument details and calibration, *J. Atmos. Oceanic Technol.*, *23*, 1397–1409, doi:10.1175/JTECH1928.1.

Weinstock, E. M., J. B. Smith, D. Sayres, J. V. Pittman, N. Allen, and J. G. Anderson (2006b), Measurements of the total water content of cirrus clouds. Part II: Instrument performance and validation, *J. Atmos. Oceanic Technol.*, *23*, 1410–1421, doi:10.1175/JTECH1929.1.

Wennberg, P. O., et al. (1994), Aircraft-borne, laser-induced fluorescence instrument for the in situ detection of hydroxyl and hydroperoxyl radicals, *Rev. Sci. Instrum.*, *65*, 1858–1876, doi:10.1063/1.1144835.

Zöger, M., et al. (1999), Fast in situ stratospheric hygrometers: A new family of balloon-borne and airborne Lyman (photofragment fluorescence hygrometers, *J. Geophys. Res.*, *104*, 1807–1816, doi:10.1029/1998JD100025.

J. G. Anderson, M. R. Sargent, D. S. Sayres, J. B. Smith, and E. M. Weinstock, Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford St., Cambridge, MA 02138, USA. (elliot@huarp.harvard.edu)

T. F. Hanisco, Earth Sciences Division, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.

E. J. Hintsa and J. R. Spackman, Cooperative Institute for Research in Environmental Sciences, University of Colorado at Boulder, Boulder, CO 80309, USA.

E. J. Moyer, Department of the Geophysical Sciences, University of Chicago, Chicago, IL 60637, USA.

J. V. Pittman, Department of Earth and Planetary Sciences, Cambridge, MA 02138, USA.

J. M. St. Clair, Geology and Planetary Sciences Division, California Institute of Technology, Pasadena, CA 91125, USA.